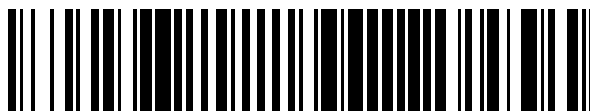


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 402 947**

51 Int. Cl.:

C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **09.04.1998** **E 09003774 (8)**

97 Fecha y número de publicación de la concesión europea: **24.10.2012** **EP 2060630**

54 Título: **PCA3, genes PCA3 y métodos de uso**

30 Prioridad:

10.04.1997 US 41836 P

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

10.05.2013

73 Titular/es:

**STICHTING KATHOLIEKE UNIVERSITEIT
UNIVERSITY MEDICAL CENTRE NIJMEGEN
(50.0%)**

GEERT GROOTEPLEIN 10

6525 GA NIJMEGEN, NL y

THE JOHNS HOPKINS UNIVERSITY (50.0%)

72 Inventor/es:

**BUSSEMAKERS, MARION y
ISAACS, WILLIAM B.**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 402 947 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

PCA3, genes PCA3 y métodos de uso

5 Campo de la invención

La presente invención se refiere, en general, a un antígeno de cáncer de próstata, PCA3. En particular, la presente invención se refiere a un método de detección de ácidos nucleicos que codifican polipéptidos o proteínas PCA3 en una muestra; sondas de ácido nucleico para la detección de ácidos nucleicos que codifican proteínas PCA3. Además, la presente divulgación se refiere a moléculas de ácido nucleico que codifican la proteína PCA3; polipéptidos y proteínas PCA3 purificados; moléculas de ácido nucleico recombinante; células que contienen las moléculas de ácido nucleico recombinante; anticuerpos que tienen afinidad de unión específicamente a polipéptidos y proteínas PCA3; hibridomas que contienen los anticuerpos; kits que contienen sondas de ácido nucleico o anticuerpos; bioensayos que usan la secuencia de ácidos nucleicos, la proteína o los anticuerpos de la presente invención para diagnosticar, evaluar o realizar un pronóstico de un mamífero afectado de cáncer de próstata; usos terapéuticos; y métodos de prevención del cáncer de próstata en un animal.

Antecedentes de la invención

El cáncer de próstata es la neoplasia maligna diagnosticada más frecuentemente y la segunda causa de muerte relacionada con el cáncer en la población masculina de los países industrializados. Cuando este carcinoma se ha extendido localmente o a zonas alejadas, no se puede ofrecer ningún tratamiento curativo. Por lo tanto, los esfuerzos para controlar la enfermedad (es decir, para reducir la mortalidad por cáncer de próstata) se han centrado en aumentar la detección del cáncer mientras todavía se encuentra confinado localmente y es potencialmente curable. Estudios dirigidos a la detección temprana del cáncer de próstata han demostrado un aumento notable en la detección de cánceres de próstata confinados en el órgano, potencialmente curables. Sin embargo, todavía no se ha hecho para tasas de mortalidad específicas de cáncer. Por otro lado, tampoco existen pruebas de que el diagnóstico temprano reduzca las tasas de mortalidad. Tanto en Estados Unidos como en Europa, siguen en curso análisis sobre la eficacia y la aceptabilidad de los programas de cribado, el problema del sobrediagnóstico y el sobretratamiento y las posibilidades de que el tratamiento temprano conduzca a la disminución de la mortalidad y la morbilidad por cáncer de próstata, y hacen de la detección temprana del cáncer de próstata un asunto controvertido (Schröder, Urology 46: 6270 (1995)).

Las medidas de las concentraciones séricas de enzimas marcadoras prostáticas tienen un valor reconocido en la detección clínica, el diagnóstico y el tratamiento del cáncer de próstata. Las dos enzimas marcadoras prostáticas de uso más extendido son la fosfatasa ácida prostática (PAP) y el antígeno específico de próstata (PSA). Normalmente, ambas enzimas se secretan desde las células epiteliales prostáticas al semen, pero en pacientes con enfermedad prostática se filtran a la circulación, donde se pueden detectar por medio de ensayos inmunológicos (Armbruster, Clin. Che. 39: 181-95 (1993)).

Además, se ha descrito que el DD3 es un gen específico de cáncer de próstata y se ha indicado que el DD3 es un marcador prometedor para el diagnóstico temprano del cáncer de próstata y que proporciona una herramienta potente para el desarrollo de nuevas estrategias de tratamiento para pacientes con cáncer de próstata (Bussemakers MJG et al.; 87ª reunión anual de la Asociación Estadounidense de Investigación Oncológica, Washington, D.C., EE. UU., 20-24 de abril de 1996. Bussemakers MJG et al., 12º Congreso de la Sociedad Europea de Oncología Urológica y Endocrinología, Essen, Alemania, 6-8 de marzo de 1997).

La fosfatasa ácida prostática, uno de los primeros marcadores séricos para la próstata, tiene una función todavía no determinada y es uno de los componentes proteicos más predominantes en las secreciones prostáticas humanas. El uso de la PAP como marcador para tumores prostáticos se complica por las similitudes estructurales de las que se ha informado entre la fosfatasa ácida específica de próstata y la fosfatasa ácida lisosómica presente en todos los tejidos. Además, existe una tendencia hacia niveles más bajos de proteína y ARNm de PAP en cáncer de próstata en comparación con la hiperplasia prostática benigna (HPB). En los últimos años, se sustituyeron las medidas de PAP por medidas de PSA sérico en el tratamiento clínico del cáncer de próstata.

El antígeno específico de próstata se identificó por varios grupos en los años 70 como una proteína específica de próstata del semen. En 1979, se purificó como un antígeno a partir de tejido de cáncer de próstata. La investigación adicional mostró que el PSA los producen exclusivamente las células epiteliales columnares de la próstata y las glándulas periuretrales. El epitelio prostático normal y el tejido hiperplásico benigno producen de hecho más proteína y ARNm de PSA de lo que lo hace el tejido de cáncer de próstata. Además, se demostró que la pérdida de diferenciación de carcinomas prostáticos se asocia con una disminución del nivel de PSA intraprostático.

Las anomalías en la arquitectura de la próstata que se producen como consecuencia de enfermedades prostáticas conducen a un aumento de la filtración de PSA (y PAP) al suero y hacen de las medidas de PSA sérico un marcador para el cáncer de próstata. A pesar del hecho de que estudios tempranos han indicado que las pruebas de PSA de diagnóstico estarían obstaculizadas por el hecho de que carecían de especificidad en la diferenciación entre HPB y

cáncer de próstata, las pruebas de PSA se introdujeron en 1986 y revolucionaron el tratamiento de pacientes con cáncer de próstata. El aumento del conocimiento de la especificidad de órgano del PSA y la relación de niveles séricos elevados de PSA y enfermedades de la próstata, así como una mejora de las técnicas de biopsia y evaluación histológica, condujeron a una apreciación del valor clínico de las pruebas de PSA, una utilidad que no ha alcanzado aún ningún otro marcador tumoral (de próstata). La clonación del gen que codifica el PSA reveló que es un miembro de la familia de genes de calicreína humanos y dio lugar al desarrollo de nuevos enfoques para el uso de PSA como marcador: se usa el método de la reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR), muy sensible, para detectar cantidades extremadamente pequeñas de células de próstata malignas en muestras de sangre de pacientes con cáncer de próstata y podría proporcionar una herramienta sensible para identificar pacientes con enfermedad micrometastásica (Moreno et al., Cancer Res. 52: 6110-12 (1992); y Katz et al., Urology 43: 765-75 (1994)).

Originalmente, el antígeno de membrana específico de próstata (PSM) se identificó usando un anticuerpo desarrollado inmunizando ratones con la fracción de membrana de células de adenocarcinoma prostático humano LNCaP. Como la PAP y el PSA, el PSM se puede detectar en próstata normal, HPB y cáncer de próstata y está ausente en la mayoría del resto de tejidos. También se han desarrollado estudios de RT-PCR para PSM para detectar células de cáncer de próstata en circulación, aunque se necesitan investigaciones adicionales para establecer la utilidad del PSM como marcador para el cáncer prostático.

En resumen, actualmente se reconoce el PSA como el principal marcador para cáncer prostático, siendo útil para cribar poblaciones seleccionadas de pacientes con síntomas que indiquen cáncer de próstata y para realizar un seguimiento de los pacientes después del tratamiento, especialmente después de la prostatectomía quirúrgica (niveles de PSA medibles indican enfermedad residual o metástasis y concentraciones crecientes de PSA indican enfermedad en recidiva). Las debilidades significativas del PSA como marcador tumoral son que (1) el PSA no siempre puede distinguir el cáncer de próstata de la HPB; y (2) que su expresión disminuye con la pérdida de diferenciación de carcinomas.

En vista del hecho de que el cáncer de próstata avanzado sigue siendo una enfermedad potencialmente mortal que afecta a una proporción muy significativa de la población masculina, sigue existiendo la necesidad de desarrollar nuevas modalidades de diagnóstico y tratamiento para cáncer de próstata (en fase tardía).

Además, se han proporcionado métodos para detectar metástasis de carcinoma en tejidos o fluidos corporales seleccionados. En particular, se han identificado secuencias de ácidos nucleicos asociadas con carcinoma diana para la detección de enfermedad residual mínima en carcinoma de pulmón (en el documento EP 0520794A (Hoffmann-La Roche; Univ California (EE. UU.)) 30 de diciembre de 1992).

La presente invención busca satisfacer estas y otras necesidades.

Sumario de la invención

La invención se refiere, en general, a moléculas de ácido nucleico aisladas que codifican PCA3 o fragmentos de la misma.

También se divulgan ácidos nucleicos para la detección específica de la presencia de ácidos nucleicos que codifican polipéptidos o proteínas PCA3 en una muestra.

La invención proporciona además un método de detección de ácido nucleico que codifica PCA3 en una muestra.

La presente invención se refiere a las realizaciones caracterizadas en las reivindicaciones. Así, se refiere a los siguientes puntos:

1. Un método para determinar si un ácido nucleico de PCA3 está presente en una muestra de prueba que comprende ARN, mostrando dicho ácido nucleico de PCA3 sobreexpresión en células de cáncer de próstata en relación con células de próstata normales o tejidos de próstata benignos, comprendiendo dicho método las etapas de:

(a) sintetizar ADNc usando ARN de la muestra de prueba como molde; y

(b) detectar entre el ADNc sintetizado un ácido nucleico que comprende un marco de lectura abierto (ORF) que codifica la SEQ ID NO: 7, en el que la Arg23 del ORF está codificada por el codón AGA que resulta del ajuste del ARN del exón 3 del PCA3 al exón 4a del PCA3, determinando de este modo que el ácido nucleico de PCA3 está presente en la muestra de prueba.

2. El método del punto 1, en el que la etapa (a) comprende alinear un cebador sintético adyacente a una secuencia diana contenida dentro del ARN de la muestra de prueba, en el que dicho cebador sintético es complementario a una secuencia contenida dentro de un ácido nucleico que comprende el exón 3 del PCA3 y el exón 4a del PCA3.

3. El método del punto 1 o 2, en el que el ácido nucleico de PCA3 comprende la secuencia del ORF que se extiende desde la posición de base 401 a la 553 de la SEQ ID NO: 6.
- 5 4. El método de uno cualquiera de los puntos 1 a 3, en el que la muestra de prueba es un extracto de ácido nucleico de células o fluidos biológicos.
5. El método de uno cualquiera de los puntos 1 a 3, en el que la muestra de prueba es una muestra biológica seleccionada del grupo que consiste en un organismo completo, una muestra de órgano, una muestra de tejido, una muestra de sangre y una muestra de células.
- 10 6. El método de uno cualquiera de los puntos 2 a 5, en el que el cebador sintético tiene una longitud de:
 - (i) desde 10 hasta 50 nucleótidos;
 - 15 (ii) desde 10 hasta 35 nucleótidos; o
 - (iii) desde 20 hasta 35 nucleótidos.
- 20 7. El método de uno cualquiera de los puntos 1 a 6, en el que dicho cebador sintético y ARN de la muestra de prueba pueden formar una estructura bicatenaria en una solución de formamida al 50 %, SSC 5x, solución de Denhardt 5x, SDS al 1 % y 100 µg/ml de ADN de esperma de salmón desnaturalizado a 68 °C.
- 25 8. El método del punto 6, en el que el cebador sintético es complementario a una secuencia contenida dentro del exón 4a.
9. El método del punto 6, en el que el cebador sintético comprende una secuencia contenida dentro del exón 3.
- 30 10. El método de uno cualquiera de los puntos 2 a 9, en el que la etapa (a) de alineación comprende combinar ARN de la muestra de prueba con el cebador sintético en una mezcla de reacción para realizar la RT-PCR.
11. El método de uno cualquiera de los puntos 1 a 10, que comprende además una etapa para determinar si el ácido nucleico de PCA3 que se determinó que estaba presente en la muestra de prueba está presente en una cantidad mayor que la cantidad presente en una muestra que comprende ácidos nucleicos aislados a partir de células de tejido de próstata normal.
- 35 12. El método del punto 11, que comprende además una etapa para detectar un ácido nucleico que codifica el PSA (antígeno específico de próstata) o un ácido nucleico que codifica el PSM (antígeno de membrana específico de próstata).
- 40 13. El método de uno cualquiera de los puntos 1 a 12, en el que la etapa (b) de detección comprende detectar la secuencia de ácidos nucleicos que codifican el ORF de PCA3 completo.
- 45 También se divulga un kit para detectar la presencia de ácido nucleico que codifica PCA3 en una muestra.
- 50 También se divulgan polipéptidos que codifican la PCA3 purificados o una porción proteica de unión a antígeno de la misma.
- 55 También se divulga una molécula de ácido nucleico recombinante que comprende, de 5' a 3', un promotor eficaz para iniciar la transcripción en una célula huésped y la molécula de ácido nucleico aislada descrita anteriormente.
- 60 También se divulga una molécula de ácido nucleico recombinante que comprende un vector y la molécula de ácido nucleico descrita anteriormente.
- 65 También se divulga una molécula de ácido nucleico de PCA3 antisentido.
- También se divulga una célula que contiene la molécula de ácido nucleico recombinante descrita anteriormente.
- También se divulga un organismo no humano que contiene la molécula de ácido nucleico recombinante descrita anteriormente.
- También se divulga un anticuerpo que tiene afinidad de unión específicamente a PCA3 o una porción portadora de epítipo de la misma.
- También se divulga un método de detección de PCA3 en una muestra.

También se divulga un método de medida de la cantidad de PCA3 en una muestra.

También se divulgan reactivos inmunógenos para inducir la protección contra células de cáncer de próstata que expresan PCA3. Preferentemente, estos reactivos inmunógenos son polipéptidos que codifican PCA3, una porción antigénica de la misma, proteínas de fusión que codifican PCA3 o proteínas de fusión que codifican porciones antigénicas de PCA3. En una realización de este tipo, estos reactivos inmunógenos funcionarían como agentes de vacuna.

También se divulga un método de detección de anticuerpos que tienen afinidad de unión específicamente a PCA3.

También se divulga un kit de diagnóstico que comprende un primer medio de recipiente que contiene el anticuerpo descrito anteriormente y un segundo medio de recipiente que contiene un conjugado que comprende un compañero de unión del anticuerpo monoclonal y una marca.

También se divulga un hibridoma que produce el anticuerpo monoclonal descrito anteriormente.

También se divulgan métodos de diagnóstico para enfermedades humanas, en particular, cáncer de próstata. Preferentemente, en el presente documento se proporciona un método de diagnóstico de la presencia o predisposición a desarrollar cáncer de próstata en un paciente.

También se divulgan métodos para usos terapéuticos que implican la totalidad o parte de (1) una secuencia de ácidos nucleicos que codifica PCA3, (2) moléculas de ácido nucleico de PCA3 antisentido, (2) proteína PCA3 o (4) anticuerpos de PCA3.

Otros objetivos y ventajas de la presente invención resultarán evidentes a partir de la siguiente descripción.

Definiciones

En la descripción siguiente, se utilizan una serie de términos de uso extendido en tecnología ADN recombinante (ADNr). Las definiciones siguientes se proporcionan con el fin de proporcionar una comprensión clara y consistente de la memoria descriptiva y las reivindicaciones, incluido el alcance que se da a tales términos.

Molécula de ácido nucleico aislada. Una "molécula de ácido nucleico aislada", como se entiende en general y se usa en el presente documento, se refiere a un polímero de nucleótidos e incluye, pero no debería limitarse a, ADN y ARN. La molécula de ácido nucleico "aislada" se purifica a partir de su estado natural *in vivo*.

ADN recombinante. Cualquier molécula de ADN formada uniendo segmentos de ADN de diferentes fuentes y producida usando tecnología de ADN recombinante (también conocida como ingeniería genética molecular).

Segmento de ADN. Un segmento de ADN, como se entiende en general y se usa en el presente documento, se refiere a una molécula que comprende un tramo lineal de nucleótidos en el que los nucleótidos están presentes en una secuencia que puede codificar, a través del código genético, una molécula que comprende una secuencia lineal de residuos de aminoácido que se denomina proteína, fragmento de proteína o polipéptido.

Gen. Una secuencia de ADN relacionada con una única cadena polipeptídica o proteína y que, como se usa en el presente documento, incluye los extremos 5' y 3' no traducidos. EL polipéptido puede estar codificado por una secuencia de longitud completa o cualquier porción de la secuencia codificante, siempre que se mantenga la actividad funcional de la proteína.

ADN complementario (ADNc). Moléculas de ácido nucleico recombinante sintetizadas por transcripción inversa de ARN mensajero ("ARNm").

Gen estructural. Una secuencia de ADN que se transcribe en un ARNm que después se traduce en una secuencia de aminoácidos características de un polipéptido específico.

Endonucleasa de restricción. Una endonucleasa de restricción (también enzima de restricción) es una enzima que tiene la capacidad de reconocer una secuencia de bases específica (habitualmente de 4, 5 o 6 pares de bases de longitud) en una molécula de ADN y de escindir la molécula de ADN en todos los lugares donde aparece esta secuencia. Por ejemplo, EcoRI reconoce la secuencia de bases GAATTC/CTTAAG.

Fragmento de restricción. Las moléculas de ADN producidas por digestión con una endonucleasa de restricción se denominan fragmentos de restricción. Cualquier genoma dado se puede digerir por una endonucleasa de restricción en particular en un conjunto discreto de fragmentos de restricción.

Electroforesis en gel de agarosa. Para detectar un polimorfismo en la longitud de los fragmentos de restricción es necesario un método analítico para fraccionar moléculas de ADN bicatenario basándose en su tamaño. La técnica

usada más comúnmente (aunque no la única) para lograr este fraccionamiento es la electroforesis en gel de agarosa. El principio de este método es que las moléculas de ADN migran a través de gel como si fuera un tamiz que retarda el movimiento de las moléculas más grandes en la mayor medida y el movimiento de las más pequeñas en la menor medida. Obsérvese que cuanto más pequeño sea el fragmento de ADN, mayor será su movilidad bajo electroforesis en el gel de agarosa.

Los fragmentos de ADN fraccionados por electroforesis en gel de agarosa se pueden visualizar directamente mediante un procedimiento de tinción si el número de fragmentos incluidos en el patrón es pequeño. Se pueden visualizar con éxito los fragmentos de ADN de genomas. No obstante, la mayoría de los genomas, incluido el genoma humano, contienen demasiadas secuencias de ADN como para producir un patrón sencillo de fragmentos de restricción. Por ejemplo, el genoma humano se digiere en aproximadamente 1.000.000 de fragmentos de ADN diferentes por *EcoRI*. Con el fin de visualizar un subconjunto pequeño de estos fragmentos, se puede aplicar una metodología a la que se hace referencia como el procedimiento de hibridación de Southern.

Procedimiento de transferencia de Southern. El propósito del procedimiento de transferencia de Southern (también denominado transferencia de bandas) es transferir físicamente ADN fraccionado por electroforesis en gel de agarosa a un papel de filtro de nitrocelulosa u otra superficie o método adecuado, al mismo tiempo que se mantienen las posiciones relativas de los fragmentos de ADN resultantes del procedimiento de fraccionamiento. La metodología usada para llevar a cabo la transferencia desde el gel de agarosa a la nitrocelulosa implica extraer el ADN del gel al papel de nitrocelulosa por acción capilar.

Hibridación de ácidos nucleicos. La hibridación de ácidos nucleicos depende del principio de que dos moléculas de ácido nucleico monocatenarias que tienen secuencias de bases complementarias volverán a formar la estructura bicatenaria termodinámicamente favorecida si se mezclan en las condiciones apropiadas. Se formará la estructura bicatenaria entre dos ácidos nucleicos monocatenarios incluso si uno está inmovilizado sobre un filtro de nitrocelulosa. En el procedimiento de hibridación de Southern, se produce esta última situación. Como se indica anteriormente, el ADN del individuo que se va a someter a prueba se digiere con una endonucleasa de restricción, se fracciona por electroforesis en gel de agarosa, se convierte a la forma monocatenaria y se transfiere a papel de nitrocelulosa, dejándola disponible para volver a alinearla con la sonda de hibridación. Se pueden encontrar ejemplos de condiciones de hibridación en Ausubel, F.M. et al., *Current protocols in Molecular Biology*, John Wiley & Sons, Inc., Nueva York, NY (1989). Se incuba un filtro de nitrocelulosa durante la noche a 68 °C con sonda marcada en una solución que contiene formamida al 50 %, alto contenido en sal (SSC 5x [20x: NaCl 3M/citrato trisódico 0,3M] o SSPE 5x [20x: NaCl 3,6M/NaH₂PO₄ 0,2M/EDTA 0,02M, a pH 7,7]), solución de Denhardt 5x, SDS al 1 % y 100 µg/ml de ADN de esperma de salmón desnaturalizado. Esto va seguido de varios lavados en SSC 0,2x/SDS al 0,1 % a una temperatura seleccionada basándose en la rigurosidad deseada: temperatura ambiente (rigurosidad baja), 42 °C (rigurosidad moderada) o 68 °C (rigurosidad alta). La temperatura seleccionada se determina basándose en la temperatura de fusión (T_f) del híbrido de ADN.

Sonda de hibridación. Para visualizar una secuencia de ADN en particular en el procedimiento de hibridación de Southern, se hace reaccionar una molécula de ADN marcada o sonda de hibridación con el ADN fraccionado unido al filtro de nitrocelulosa. Las propias zonas del filtro portadoras de secuencias de ADN complementarias a la sonda de ADN marcada se marcan como consecuencia de la reacción de realineación. Se visualizan las zonas del filtro que muestran un marcaje de este tipo. En general, la sonda de hibridación se produce por clonación molecular de una secuencia de ADN específica.

Oligonucleótido u oligómero. Una molécula que comprende dos o más desoxirribonucleótidos o ribonucleótidos, preferentemente más de tres. Su tamaño exacto dependerá de muchos factores, que a su vez dependen de la función última o el uso del oligonucleótido. Un oligonucleótido se puede derivar sintéticamente o por clonación.

Amplificación de secuencias. Un método para generar grandes cantidades de secuencia diana. En general, se alinean uno o más cebadores de amplificación con una secuencia de ácidos nucleicos. Usando enzimas apropiadas, se amplifican secuencias que se encuentran adyacentes a, o entre, los cebadores.

Cebador de amplificación. Un oligonucleótido que se puede alinear adyacente a una secuencia diana y servir como punto de inicio para la síntesis de ADN cuando se pone en condiciones en las cuales se inicia la síntesis de un producto de extensión de cebador que es complementario a una hebra de ácido nucleico complementaria.

Molécula de ácido nucleico antisentido. Una "molécula de ácido nucleico antisentido" se refiere en el presente documento a una molécula que puede formar un dúplex o tríplex estable con una porción de su secuencia de ácidos nucleicos diana (ADN o ARN). El uso de moléculas de ácido nucleico antisentido y el diseño y la modificación de tales moléculas son bien conocidos en la técnica como se describe, por ejemplo, en los documentos WO 96/32966, WO 96/11266, WO 94/15646, WO 93/08845 y USP 5.593.974. Las moléculas de ácido nucleico de acuerdo con la presente invención se pueden derivar de las secuencias de ácido nucleico de la presente invención y modificar de acuerdo con métodos bien conocidos. Por ejemplo, algunas moléculas antisentido se pueden diseñar para que sean más resistentes a la degradación, para aumentar su afinidad por su secuencia diana, para afectar a su transporte hacia tipos celulares o compartimentos celulares escogidos y/o para potenciar su solubilidad en lípidos usando

análogos de nucleótidos y/o sustituyendo fragmentos químicos escogidos de los mismos, como se conoce comúnmente en la técnica.

Vector. Un ADN plasmídico o de fago u otra secuencia de ADN en la cual se puede insertar ADN para clonarlo. El vector se puede replicar de forma autónoma en una célula huésped y se puede caracterizar adicionalmente por uno o un número menor de sitios de reconocimiento de endonucleasa en los que se pueden cortar estas secuencias de ADN de una manera determinable y en los que se puede insertar ADN. El vector puede contener además un marcador adecuado para su uso en la identificación de células transformadas con el vector. Son marcadores, por ejemplo, la resistencia a tetraciclina o la resistencia a ampicilina. En ocasiones, se usan las palabras "vehículo de clonación" para "vector".

Expresión. La expresión es el proceso por el que un gen estructural produce un polipéptido. Implica la transcripción del gen en ARNm y la traducción de este ARNm en polipéptido(s).

Vector de expresión. Un vector o vehículo similar a un vector de clonación pero que puede expresar un gen que se ha clonado en él, después de la transformación en un huésped. Habitualmente, el gen clonado se sitúa bajo el control de (es decir, unido de forma funcional a) determinadas secuencias de control tales como secuencias promotoras.

Las secuencias de control de la expresión varían en función de si el vector está diseñado para expresar el gen unido de forma funcional en un huésped procariota o eucariota y pueden contener adicionalmente elementos transcripcionales tales como elementos potenciadores, secuencias de terminación, elementos específicos de tejido y/o sitios de iniciación traduccional y de terminación traduccional.

Derivado funcional. Un "derivado funcional" de una secuencia, bien una proteína o un ácido nucleico, es una molécula que posee una actividad biológica (bien funcional o estructural) que es sustancialmente similar a una actividad biológica de la proteína o la secuencia de ácidos nucleicos. Un derivado funcional de una proteína puede contener modificaciones postraduccionales tales como carbohidratos unidos covalentemente, dependiendo de la necesidad de dichas modificaciones para la realización de una función específica. Se pretende que el término "derivado funcional" incluya los "fragmentos", "segmentos", "variantes", "análogos" o "derivados químicos" de una molécula.

Como se usa en el presente documento, se dice que una molécula es un "derivado químico" de otra molécula cuando contiene restos químicos adicionales que normalmente no forman parte de la molécula. Estos restos pueden mejorar la solubilidad, la absorción, la semivida biológica y similares de la molécula. De forma alternativa, los restos pueden reducir la toxicidad de la molécula, eliminar o atenuar cualquier efecto secundario no deseable de la molécula, y similares. En Remington's Pharmaceutical Sciences (1980) se divulgan restos capaces de mediar efectos de este tipo. En la técnica, se conocen bien procedimientos para acoplar estos restos a una molécula.

Variante. Se pretende que una "variante" de una proteína o ácido nucleico se refiera a una molécula sustancialmente similar en estructura y actividad biológica bien a la proteína o al ácido nucleico. Por tanto, con la condición de que dos moléculas posean una actividad común y puedan sustituirse entres sí, se consideran variantes como se usa ese término en el presente documento, incluso si la composición o la estructura secundaria, terciaria o cuaternaria de una de las moléculas no es idéntica a la que se encuentra en la otra, o si la secuencia de aminoácidos o nucleótidos no es idéntica.

Alelo. Un "alelo" es una forma alternativa de un gen que ocupa un locus dado en el cromosoma.

Mutación. Una "mutación" es cualquier cambio detectable en el material genético que se puede transmitir a las células hijas incluso a generaciones sucesivas, dando lugar a células mutantes o individuos mutantes. Si los descendientes de una célula mutante sólo dan lugar a células somáticas en organismos pluricelulares, aparece una mancha o zona de células mutante. Las mutaciones en la línea germinal de organismos que se reproducen sexualmente se pueden transmitir por los gametos a la siguiente generación, dando lugar a un individuo con la nueva condición mutante tanto en sus células somáticas como en las germinales. Una mutación puede ser cualquier cambio (o una combinación) detectable, no natural, que afecta a la constitución química o física, mutabilidad, replicación, función fenotípica o recombinación de uno o más desoxirribonucleótidos; se pueden añadir nucleótidos, deletarlos, sustituirlos por, invertirlos o transponerlos a posiciones nuevas con o sin inversión. Las mutaciones se producen espontáneamente y se pueden inducir de forma experimental por aplicación de mutágenos. Una variación mutante de una molécula de ácido nucleico resulta de una mutación. Un polipéptido mutante puede resultar de una molécula de ácido nucleico mutante.

Especie. Una "especie" es un grupo de poblaciones naturales que se entrecruzan realmente o potencialmente. Una variación de especie dentro de una molécula de ácido nucleico o proteína es un cambio en la secuencia de ácidos nucleicos o aminoácidos que se produce entre especies y se puede determinar por secuenciación de ADN de la molécula en cuestión.

Electroforesis en gel de poliacrilamida (PAGE). La técnica de uso más extendido (aunque no la única) para lograr un fraccionamiento de polipéptidos basado en el tamaño es la electroforesis en gel de poliacrilamida. El principio de este método es que las moléculas de polipéptido migran a través de gel como si fuera un tamiz que retarda el movimiento de las moléculas más grandes en la mayor medida y el movimiento de las más pequeñas en la menor medida. Obsérvese que cuanto más pequeño sea el fragmento de polipéptido, mayor será su movilidad bajo electroforesis en el gel de poliacrilamida. Normalmente, tanto antes como durante la electroforesis, los polipéptidos están expuestos continuamente al detergente dodecilsulfato de sodio (SDS), condiciones bajo las cuales los polipéptidos se desnaturalizan. Los geles nativos se realizaron en ausencia de SDS. Los polipéptidos fraccionados por electroforesis en gel de poliacrilamida se pueden visualizar directamente mediante un procedimiento de tinción si el número de componentes polipeptídicos es pequeño.

Procedimiento de transferencia western. El propósito del procedimiento de transferencia western (también denominado transferencia de bandas) es transferir físicamente polipéptidos fraccionados por electroforesis en gel de poliacrilamida a un papel de filtro de nitrocelulosa u otra superficie o método adecuado, al mismo tiempo que se mantienen las posiciones relativas de los polipéptidos resultantes del procedimiento de fraccionamiento. Después, se sondan las bandas con un anticuerpo que se une específicamente al polipéptido de interés.

Purificado. Una proteína o un ácido nucleico "purificado" es una proteína o un ácido nucleico que se ha separado de un componente celular. Las proteínas o los ácidos nucleicos "purificados" se han purificado hasta un nivel que no se encuentra en la naturaleza.

Sustancialmente puro. Una proteína o un ácido nucleico "sustancialmente puro" es una proteína o un ácido nucleico que carece de todos los demás componentes celulares.

Breve descripción de los dibujos

Habiéndose descrito por tanto la invención en general, se hará referencia ahora a los dibujos que se acompañan, que muestran a modo de ilustración una realización preferida de la misma y en los que:

La figura 1 muestra la estructura genómica del gen PCA3.

Figura 2A-B, (A) una estructura de ADNc de PCA3; (B) una secuencia de nucleótidos y de aminoácidos de ADNc de PCA3 (SEQ ID NO:1 y 2).

La figura 3A-B es una representación esquemática que compara los clones de ADNc pMB9 y λ DD3.6.

La figura 4 muestra la localización de un sitio de inicio de la transcripción (TSS) de PCA3. El sitio de inicio de la transcripción se determinó por extensión de cebador (PE), mapeo con nucleasa S1 (S1) y ensayos de amplificación rápida de extremos de ADNc 5' (RACE).

Figura 5A-B, (A) una estructura de ADNc de PCA3; (B) una secuencia de nucleótidos y de aminoácidos de ADNc de PCA3 (SEQ ID NO:6 y 7); las señales putativas de poliadenilación están subrayadas.

Otros objetivos, ventajas y características divulgados también en el presente documento resultarán más evidentes tras la lectura de la siguiente descripción no restrictiva de realizaciones preferidas con referencia a los dibujos adjuntos, que son ejemplares.

Descripción de la realización preferida

En aras de la claridad de la divulgación, y no a modo de limitación, la descripción detallada de la invención se divide en las siguientes subsecciones:

I. Moléculas de ácido nucleico aisladas que codifican polipéptidos de PCA3.

II. Polipéptidos de PCA3 purificados.

III. Una sonda de ácido nucleico para la detección específica de ácido nucleico de PCA3.

IV. Un método de detección de la presencia de ácido nucleico de PCA3 en una muestra.

V. Un kit para detectar la presencia de ácido nucleico de PCA3 en una muestra.

VI. Construcciones de ADN que comprenden una molécula de ácido nucleico de PCA3 y células que contienen estas construcciones.

VII. Un anticuerpo que tiene afinidad de unión a un polipéptido de PCA3 y un hibridoma que contiene el anticuerpo.

VIII. Un método de detección de un polipéptido o un anticuerpo de PCA3 en una muestra.

IX. Un kit de diagnóstico que comprende una proteína o un anticuerpo de PCA3

X. Cribado de diagnóstico

XI. Tratamientos terapéuticos

XII. Animales no humanos con PCA3 transgénicos

I. Moléculas de ácido nucleico aisladas que codifican polipéptidos de PCA3

En el contexto de la presente invención se divulgan moléculas de ácido nucleico de PCA3 aisladas (purificadas). Preferentemente, la molécula de ácido nucleico de PCA3 comprende una secuencia de polinucleótidos idéntica en al menos un 90 % (más preferentemente idéntica en un 95 %, 96 %, 97 %, 98 %, 99 % o 100 %) a una secuencia seleccionada del grupo que consiste en:

(a) una secuencia de nucleótidos que codifica un polipéptido de PCA3 que comprende la secuencia de aminoácidos completa de SEQ ID NO:2 o SEQ ID NO:7;

(b) una secuencia de nucleótidos que codifica un polipéptido de PCA3 que comprende la secuencia de aminoácidos completa codificada por el clon de polinucleótidos contenido en el depósito en la Centraal voor Schimmelcultures con número de acceso CBS 682.97;

(c) una secuencia de nucleótidos que codifica un polipéptido de PCA3 que comprende la secuencia de aminoácidos completa codificada por el clon de polinucleótidos contenido en el depósito en la Centraal voor Schimmelcultures con número de acceso CBS 100521; y

(d) una secuencia de nucleótidos complementaria a las secuencias de (a), (b) o (c).

El pMB9 es un clon de ADNc de PCA3 que contiene los exones 1, 2, 3, 4a y 4b del gen PCA3, el pMB9 se depositó en la Phabagen Collection, Universidad de Utrecht, Padualaan 8, 3584 CH Utrecht (que es una división del Centraalbureau voor Schimmelcultures, Oosterstratt1, Postbus 273, 3740 AG Baarn) bajo la normativa del Tratado de Budapest el 10 de abril de 1997 con número de acceso CBS 682.97.

El λ DD3.6 es un clon de ADNc de PCA3 que contiene los exones 3, 4a, 4b, 4c y 4d. El λ DD3.6 se depositó en la Phabagen Collection, Universidad de Utrecht, Padualaan 8, 3584 CH Utrecht (que es una división del Centraalbureau voor Schimmelcultures, Oosterstratt1, Postbus 273, 3740 AG Baarn) bajo la normativa del Tratado de Budapest el 27 de marzo de 1998 con número de acceso CBS 100521.

Como se divulga en el presente documento, la molécula de ácido nucleico aislada comprende una secuencia de nucleótidos de PCA3 con una identidad o similitud de más de un 90 % con la secuencia de nucleótidos presente en la SEQ ID NO:1 (preferentemente de más de un 95 %, 96 %, 97 %, 98 %, 99 % o 100 %). En otra realización preferida, la molécula de ácido nucleico aislada comprende la secuencia codificante de PCA3 presente en la SEQ ID NO:1. En otra realización, la molécula de ácido nucleico aislada codifica la secuencia de aminoácidos presente en la SEQ ID NO:2 o SEQ ID NO:7. En otra realización más, la molécula de ácido nucleico aislada comprende una secuencia de nucleótidos de PCA3 con una identidad o similitud de más de un 90 % con la secuencia de nucleótidos presente en la SEQ ID NO:6 (preferentemente de más de un 95 %, 96 %, 97 %, 98 %, 99 % o 100 %). En otra realización preferida, la molécula de ácido nucleico aislada comprende la secuencia codificante de PCA3 presente en la SEQ ID NO:6.

También se divulgan dentro del alcance de la presente descripción ácidos nucleicos que comprenden variantes de ajuste de ADNc de PCA3 o secuencias de polinucleótidos que son idénticas a ellas en al menos un 90 %, preferentemente idénticas a ellas en al menos un 95 %. En vista del hecho de que son posibles prácticamente todas las combinaciones de exones, los ejemplos no limitantes de tales variantes de ajuste incluyen ácidos nucleicos de PCA3 aislados que comprenden los exones 1, 2, 3, 4a y 4b (SEQ ID NO:1); los exones 1, 3, 4a, 4b y 4c (SEQ ID NO:3 y la región 4c que es contigua a la región 4b, véase la figura 1); los exones 1, 3, 4a, 4b, 4c y 4d (SEQ ID NO:3 y la región 4c que es contigua a la región 4b y la región 4d que es contigua a la región 4c, véase la figura 1); los exones 1, 3, 4a y 4b (SEQ ID NO:3); los exones 1, 3 y 4a (SEQ ID NO:4); los exones 1, 2, 3, 4a, 4b, 4c y 4d (SEQ ID NO:6). Preferentemente, la molécula de ácido nucleico de PCA3 comprende una secuencia de polinucleótidos idéntica en al menos un 90 % (más preferentemente, idéntica en al menos un 95 %, 96 %, 97 %, 98 %, 99 % o 100 %) a una o más de las variantes de ajuste descritas anteriormente.

También se divulgan los equivalentes funcionales de las moléculas de ácido nucleico aisladas descritas en el presente documento y derivados de las mismas. Por ejemplo, las secuencias de ácidos nucleicos representadas en

la SEQ ID NO:1 o la SEQ ID NO:6 pueden estar alteradas por sustituciones, adiciones o deleciones que proporcionan moléculas funcionalmente equivalentes. Debido a la degeneración de las secuencias codificantes de nucleótidos, se pueden usar otras secuencia de ADN que codifican sustancialmente la misma secuencia de aminoácidos que se representa en la SEQ ID NO:2 y la SEQ ID NO:7 en la práctica de la presente divulgación.

Estas incluyen, pero sin limitación, secuencia de nucleótidos que comprenden la totalidad o porciones del ácido nucleico de PCA3 en la SEQ ID NO:1, 3, 4 o 6, que están alteradas por la sustitución de diferentes codones que codifican un residuo de aminoácido funcionalmente equivalente dentro de la secuencia.

Además, la secuencia de ácidos nucleicos puede comprender una secuencia de nucleótidos que es consecuencia de la adición, deleción o sustitución de al menos un nucleótido en el extremo 5' y/o el extremo 3' de la fórmula de ácido nucleico mostrada en la SEQ ID NO:1, 3, 4 o 6 o un derivado de la misma. A este respecto, se puede usar cualquier nucleótido o polinucleótido, con la condición de que su adición, deleción o sustitución no altere sustancialmente la secuencia de aminoácidos de SEQ ID NO:2 o SEQ ID NO:7 que está codificada por la secuencia de nucleótidos. Además, a la molécula de ácido nucleico de la presente descripción se le pueden añadir, en caso necesario, sitios de reconocimiento por endonucleasas de restricción en el extremo 5' y/o el extremo 3'. Por lo tanto, todas las variaciones de la secuencia de nucleótidos de la secuencia de codificación de nucleótidos de PCA3 y fragmentos de la misma permitidas por el código genético se incluyen en la presente divulgación.

Además, es posible eliminar codones o sustituir uno o más codones por codones distintos de los codones degenerados para producir un polipéptido estructuralmente modificado, pero que tiene sustancialmente la misma utilidad o actividad que las del polipéptido producido por la molécula de ácido nucleico no modificada. Como se reconoce en la técnica, los dos polipéptidos son funcionalmente equivalentes, como lo son las dos moléculas de ácido nucleico que dan lugar a su producción, aunque las diferencias entre las moléculas de ácido nucleico no están relacionadas con la degeneración del código genético.

A. Aislamiento de ácido nucleico

También se divulgan que se proporcionan moléculas de ácido nucleico aisladas que codifican polipéptidos que tienen secuencias de aminoácidos que corresponden a PCA3. En particular, se puede aislar la molécula de ácido nucleico a partir de una muestra biológica que contenga ADN o ARN de PCA3.

Se puede aislar la molécula de ácido nucleico a partir de una muestra biológica que contenga ARN de PCA3 usando las técnicas de clonación de ADNc e hibridación sustractiva. También se puede aislar la molécula de ácido nucleico a partir de una colección de ADNc usando una sonda homóloga.

Se puede aislar la molécula de ácido nucleico a partir de una muestra biológica que contenga ADN genómica o a partir de una colección genómica. Las muestras biológicas adecuadas incluyen, pro sin limitación, organismos completos, órganos, tejidos, sangre y células. El método de obtención de la muestra biológica variará en función de la naturaleza de la muestra.

Un experto en la técnica se dará cuenta de que los genomas pueden estar sujetos a ligeras variaciones alélicas entre individuos. Por lo tanto, también se pretende que la molécula de ácido nucleico aislada incluya variaciones alélicas, siempre que la secuencia sea un derivado funcional de la secuencia codificante de PCA3. Cuando un alelo de PCA3 no codifica la secuencia idéntica a la que se encuentra en las SEQ ID NO:1 o 6, se puede aislar e identificar como PCA3 usando las mismas técnicas usadas en el presente documento, y especialmente técnicas de PCR, para amplificar el gen apropiado con cebadores basados en las secuencias divulgadas en el presente documento.

Un experto en la técnica se dará cuenta de que organismos que no son humanos también contendrán genes PCA3 (por ejemplo, eucariotas; más específicamente, mamíferos, aves, peces y plantas; más específicamente, gorilas, macacos rhesus y chimpancés). Se pretende que la divulgación incluya, pero sin limitación moléculas de ácido nucleico de PCA3 aisladas a partir de los organismos descritos anteriormente.

B. Síntesis de ácido nucleico

Se pretende que las moléculas de ácido nucleico como se divulgan en el presente documento también incluyan las sintetizadas químicamente. Por ejemplo, se puede diseñar una molécula de ácido nucleico con la secuencia de nucleótidos que codifica el producto de expresión de un gen PCA3, y en caso necesario dividirla en fragmentos más pequeños apropiados. Después, se puede sintetizar un oligómero que corresponda a la molécula de ácido nucleico o a cada uno de los fragmentos divididos. Estos oligonucleótidos sintéticos se pueden preparar, por ejemplo, por el método de triéster de Matteucci et al., J. Am. Chem. Soc. 103:3185-3191 (1981) o mediante el uso de un sintetizador de ADN automatizado.

Un oligonucleótido se puede derivar sintéticamente o por clonación. En caso necesario, se puede fosforilar los extremos 5' de los oligómeros usando la polinucleótido cinasa de T4. La fosforilación con cinasas de hebras individuales antes de la alineación o el marcaje se puede llevar a cabo usando un exceso de la enzima. Si la

fosforilación con cinasas es para el marcaje de una sonda, el ATP puede contener radioisótopos de actividad específica alta. Después, se puede someter el oligómero de ADN a alineación y ligación con ligasa de T4 o similar.

II. Polipéptidos de PCA3 purificados

En el contexto de la presente invención se divulga un polipéptido purificado (preferentemente, sustancialmente puro) que tiene una secuencia de aminoácidos que corresponde a PCA3, o una variante funcional del mismo. En una realización preferida, el polipéptido tiene la secuencia de aminoácidos descrita en las SEQ ID NO:2 o 7 o variantes mutantes o de especie de las mismas, o con una identidad de al menos un 80 % o una similitud de al menos un 90 % con las mismas (preferentemente, una identidad de al menos un 90 %, 95 %, 96 %, 97 %, 98 % o 99 % o una similitud de al menos un 95 %, 96 %, 97 %, 98 % o 99 % con las mismas), o al menos 6 aminoácidos contiguos de las mismas (preferentemente, al menos 10, 15, 20, 25 o 50 aminoácidos contiguos de las mismas).

También se divulgan epítopos de PCA3. El epítipo de estos polipéptidos es un epítipo antigénico o inmunógeno. Un epítipo inmunógeno es la parte de la proteína que desencadena una respuesta de anticuerpos cuando la proteína completa es el inmunógeno. Un epítipo antigénico es un fragmento de la proteína que puede desencadenar una respuesta antigénica. En la técnica se conocen bien métodos de selección de fragmentos de epítipo antigénico. Véase, Sutcliffe et al., Science 219:660-666 (1983). Los polipéptidos y péptidos portadores de epítopos antigénicos de la invención son útiles para obtener una respuesta inmunitaria que reconoce específicamente los polipéptidos. Los polipéptidos y péptidos portadores de epítopos antigénicos de la descripción comprenden al menos 7 aminoácidos (preferentemente, 9, 10, 12, 15 o 20 aminoácidos) de las proteínas de la descripción. Un ejemplo de un péptido antigénico es HTQEAQKEAQR (SEQ ID NO:5).

Se pueden preparar variantes de secuencia de aminoácidos de PCA3 por mutaciones en el ADN. Estas variantes incluyen, por ejemplo, deleciones de, o inserciones o sustituciones de, residuos del interior de la secuencia de aminoácidos que se muestra en las SEQ ID NO:2 o 7. También se puede realizar cualquier combinación de deleción, inserción y sustitución para llegar a la construcción final, con la condición de que la construcción final posea la actividad deseada.

Aunque el sitio para introducir una variación de secuencia de aminoácidos está predeterminado, la mutación *per se* no está necesariamente predeterminada. Por ejemplo, para optimizar el funcionamiento de una mutación en un sitio dado, se puede realizar mutagénesis aleatoria en la región o el codón diana y cribar las variantes de PCA3 expresadas para detectar la combinación óptima de actividad deseada. Se conocen bien técnicas para realizar mutaciones por sustitución en sitios predeterminados en ADN que tiene una secuencia conocida, por ejemplo, mutagénesis específica de sitio.

Preferentemente, la preparación de una variante de PCA3 de acuerdo con el presente documento se consigue por mutagénesis específica de sitio de ADN que codifica una versión variante o no variante de la proteína preparada anteriormente. La mutagénesis específica de sitio permite la producción de variantes de PCA3 a través del uso de secuencias de oligonucleótidos específicas que codifican la secuencia de ADN de la mutación deseada. En general, la técnica de mutagénesis específica de sitio se conoce bien en la técnica, como se ejemplifica por publicaciones tales como Adelman et al., DNA 2:183 (1983) y Ausubel et al. "Current Protocols in Molecular Biology", J. Wiley & Sons, NY, NY, 1996.

En general, las deleciones de secuencias de aminoácidos varían desde aproximadamente 1 hasta 30 residuos, más preferentemente de 1 hasta 10 residuos, y normalmente son contiguos.

Las inserciones de secuencias de aminoácidos incluyen fusiones amino y/o carboxiloterminales de desde un residuo hasta polipéptidos de longitud esencialmente no restringida, así como inserciones intrasecuencia de residuos de aminoácido individuales o múltiples. En general, las inserciones intrasecuencia (es decir, inserciones dentro de la secuencia de PCA3 completa) pueden variar desde aproximadamente 1 hasta 10 residuos, más preferentemente de 1 a 5.

El tercer grupo de variantes son aquellas en las que se ha eliminado al menos un residuo de aminoácido de la molécula de PCA3, y preferentemente sólo uno, y se ha insertado un residuo diferente en su lugar. Preferentemente, estas sustituciones se realizan de acuerdo con la siguiente tabla 1 cuando se desea modular con precisión las características de PCA3.

Tabla 1

Residuo original	Sustituciones ejemplares
Ala	gly; ser
Arg	lys
Asn	gln; his

Asp	glu
Cys	ser
Gln	asn
Glu	asp
Gly	ala; pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; tyr; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

Se realizan cambios sustanciales en la identidad funcional o inmunológica seleccionando sustituciones que son menos conservadoras que las de la tabla 1, es decir, seleccionando residuos que difieren más significativamente en su efecto de mantener (a) la estructura del esqueleto del polipéptido en la zona de la sustitución, por ejemplo, como una conformación de lámina o hélice, (b) la carga o hidrofobicidad de la molécula en el sitio diana o (c) el volumen de la cadena lateral. Las sustituciones que se esperan, en general, son aquellas en las que (a) glicina y/o prolina se sustituye por otro aminoácido o se deleciona o inserta; (b) un residuo hidrófobo, por ejemplo, serilo o treonilo, se sustituye por (o con) un residuo hidrófobo, por ejemplo, leucilo, isoleucilo, fenilalanilo, valilo o alanilo; (c) un residuo de cisteína se sustituye por (o con) cualquier otro residuo; (d) un residuo que tiene una cadena lateral electropositiva, por ejemplo, lisilo, arginilo o histidilo, se sustituye por (o con) un residuo que tiene una carga electronegativa, por ejemplo, glutamilo o aspartilo; o (e) un residuo que tiene una cadena lateral voluminosa, por ejemplo, fenilalanina, se sustituye por (o con) uno que no tiene una cadena lateral de este tipo, por ejemplo, glicina.

Se espera que algunas delecciones e inserciones no produzcan cambios radicales en las características de la PCA3. Sin embargo, cuando es difícil predecir el efecto exacto de la sustitución, delección o inserción antes de realizarlas, un experto en la técnica apreciará que el efecto se evaluará mediante ensayos de cribado rutinarios. Por ejemplo, normalmente una variante se prepara por mutagénesis específica de sitio del ácido nucleico que codifica la PCA3 nativa, la expresión del ácido nucleico variante en cultivo celular recombinante y, opcionalmente, purificación a partir del cultivo celular, por ejemplo, por adsorción de inmutafinidad en una columna (para absorber la variante uniéndola a al menos un epítipo inmunitario restante). La actividad del lisado celular o la molécula de PCA3 purificada se criba después en un ensayo de cribado adecuado para la característica deseada. Por ejemplo, se mide un cambio en el carácter inmunológico de la molécula de PCA3, tal como la afinidad por un anticuerpo dado, mediante un inmunoensayo de tipo competitivo. Se miden cambios en la actividad inmunomoduladora mediante el ensayo apropiado. Se ensayan modificaciones de propiedades de la proteína tales como estabilidad de oxidorreducción o térmica, hidrofobicidad, susceptibilidad a degradación proteolítica o tendencia a agregarse con vehículos o en multímeros, mediante métodos bien conocidos para el experto en la técnica.

Se pueden utilizar una variedad de metodologías conocidas en la técnica para obtener el péptido de la presente descripción. En una realización, se purifica el péptido a partir de tejidos o células que producen el péptido de forma natural. De forma alternativa, se pueden usar fragmentos del ácido nucleico aislado descrito anteriormente para expresar la proteína PCA3 en cualquier organismo. Las muestras de la presente invención incluyen células, extractos proteicos o extractos de membrana de células, o fluidos biológicos. La muestra variará basándose en el formato de ensayo, el método de detección y la naturaleza de los tejidos, células o extractos usados como muestra.

Se puede usar cualquier organismo como fuente para el péptido de la descripción, siempre que el organismo fuente contenga de forma natural este péptido. Como se usa en el presente documento, "organismo fuente" se refiere al organismo original a partir de cual se deriva la secuencia de aminoácidos de la subunidad, independientemente del organismo, la subunidad se expresa en una forma aislada en última instancia.

Un experto en la técnica puede seguir fácilmente métodos conocidos para aislar proteínas con el fin de obtener el péptido son contaminantes naturales. Estos incluyen, pero sin limitación: inmunocromatografía, cromatografía de

exclusión por tamaño, HPLC, cromatografía de intercambio iónico y cromatografía de inmutaoafinidad.

En una realización preferida, los procedimientos de purificación comprenden cromatografía de intercambio iónico y cromatografía de exclusión por tamaño. Se puede emplear una cualquiera de un gran número de resinas de intercambio iónico conocidas en la técnica, incluyendo, por ejemplo, monoQ, sefarsa Q, macro-prepQ, AG1-X2 o HQ. Los ejemplos de resinas de exclusión por tamaño adecuadas incluyen, pero sin limitación, Superdex 200, Superose 12 y Sephacryl 200. Se puede conseguir la elución con soluciones acuosas de cloruro de potasio o cloruro de sodio a concentraciones que varían desde 0,01 M hasta 2,0 M.

III. Un ácido nucleico para la detección específica de ácido nucleico de PCA3

En otra realización, la presente invención se refiere a un ácido nucleico para la detección específica de la presencia de ácido nucleico de PCA3 en una muestra que comprende las moléculas de ácido nucleico descritas anteriormente o al menos un fragmento de las mismas que se une en condiciones rigurosas a ácido nucleico de PCA3.

Además, también se divulga un ácido nucleico aislado que consiste en de 10 a 1000 nucleótidos (preferentemente, 10 a 500, 10 a 100, 10 a 50, 10 a 35, 20 a 1000, 20 a 500, 20 a 100, 20 a 50 o 20 a 35) que hibrida de forma preferente con ARN o ADN que codifica PCA3 o con un gen PCA3, pero no con ARN o ADN que no está relacionado con PCA3, en el que dicha sonda de ácido nucleico es o es complementaria a una secuencia de nucleótidos que consiste en el menos 10 nucleótidos consecutivos (preferentemente, 15, 18, 20, 25 o 30) de la molécula de ácido nucleico que comprende una secuencia de polinucleótidos idéntica en al menos un 90 % a una secuencia seleccionada del grupo que consiste en:

(a) una secuencia de nucleótidos que codifica el polipéptido de PCA3 que comprende la secuencia de aminoácidos completa de SEQ ID NO:2 o 7

(b) una secuencia de nucleótidos que codifica el polipéptido de PCA3 que comprende la secuencia de aminoácidos completa codificada por el clon de polinucleótidos contenido en el depósito en la Centraal voor Schimmelcultures con número de acceso CBS 682.97, respectivamente;

(c) una secuencia de nucleótidos que codifica el polipéptido de PCA3 que comprende la secuencia de aminoácidos completa codificada por el clon de polinucleótidos contenido en el depósito en la Centraal voor Schimmelcultures con número de acceso CBS 100521, respectivamente;

(d) una secuencia de nucleótidos que codifica el gen PCA3 que comprende la secuencia de nucleótidos completa de SEQ ID NO:1, 3, 4 o 6

(e) una secuencia de nucleótidos que codifica un exón de gen PCA3 que comprende los nucleótidos 1-98, 99-263, 264-446, 447-985 o 986-2037 como se describe en la SEQ ID NO:1;

(f) una secuencia de nucleótidos que codifica un exón de gen PCA3 que comprende los nucleótidos 1-120, 121-285, 286-468, 469-1007, 1008-2066, 2067-2622 o 2623-3582 como se describe en la SEQ ID NO:6;

(g) una secuencia de nucleótidos complementaria a cualquiera de las secuencias de (a), (b), (c), (d), (e) o (f).

(h) una secuencia de nucleótidos como previamente se describe más arriba.

Preferentemente, el ácido nucleico no hibrida específicamente con los nucleótidos 511-985 de la SEQ ID NO:1, con los nucleótidos 567-961 de la SEQ ID NO:1, con los nucleótidos 533-1007 de la SEQ ID NO:6 o con los nucleótidos 589-983 de la SEQ ID NO:6.

Las secuencias complementarias también se conocen como ácidos nucleicos antisentido cuando comprenden secuencias que son complementarias a la hebra codificante.

En la tabla 2, a continuación, se describen ejemplos de sondas de ácido nucleico específicas que se pueden usar en la presente invención.

Tabla 2: sondas de ácido nucleico

	Tamaño (n.º de bases)	Nucleótidos
Exón 1	98	1-98 de SEQ ID NO:1
Exón 2	165	99-263 de SEQ ID NO:1
Exón 3	183	264-446 de SEQ ID NO:1
Exón 4a	539	447-985 de SEQ ID NO:1

Exón 4b	1052	986-2037 de SEQ ID NO:1
Sonda 1	20	1-20 de SEQ ID NO:1
Sonda 2	30	1-30 de SEQ ID NO:1
Sonda 3	40	1-40 de SEQ ID NO:1
Sonda 4	20	381-400 de SEQ ID NO:1
Sonda 5	30	381-410 de SEQ ID NO:1
Sonda 6	20	401-420 de SEQ ID NO:1
Sonda 7	30	401-430 de SEQ ID NO:1
Sonda 8	20	511-530 de SEQ ID NO:1
Sonda 9	30	501-530 de SEQ ID NO:1
Sonda 10	20	77-98 de SEQ ID NO:1
Sonda 11	20	99-118 de SEQ ID NO:1
Sonda 12	20	244-263 de SEQ ID NO:1
Sonda 13	20	264-283 de SEQ ID NO:1
Sonda 14	20	427-446 de SEQ ID NO:1
Sonda 15	20	447-466 de SEQ ID NO:1
Exón 1	120	1-120 de SEQ ID NO:6
Exón 2	165	121-285 de SEQ ID NO:6
Exón 3	183	286-468 de SEQ ID NO:6
Exón 4a	539	469-1007 de SEQ ID NO:6
Exón 4b	1059	1008-2066 de SEQ ID NO:6
Exón 4c	556	2067-2622 de SEQ ID NO:6
Exón 4d	960	2623-3582 de SEQ ID NO:6

Por supuesto, como se entenderá por el experto en la técnica, se pueden diseñar una multitud de sondas adicionales de la misma o de otra región de SEQ ID NO:1 así como de SEQ ID NO:6 y otras secuencias de la presente descripción.

5 Se puede usar la sonda de ácido nucleico para sondear una colección de ADNc o cromosómica apropiada por métodos de hibridación normales para obtener otra molécula de ácido nucleico de la presente descripción. Se puede preparar una colección de ADNc o ADN cromosómico a partir de células apropiadas de acuerdo con métodos reconocidos en la técnica (cf. Molecular Cloning: A Laboratory Manual, segunda edición, editado por Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, 1989).

10 Como alternativa, se lleva a cabo la síntesis química para obtener sondas de ácido nucleico que tienen secuencias de nucleótidos que se corresponden con porciones N terminales y C terminales de la secuencia de aminoácidos de PCA3. Por tanto, se pueden usar las sondas de ácido nucleico sintetizadas como cebadores en una reacción en cadena de la polimerasa (PCR) llevada a cabo de acuerdo con técnicas de PCR reconocidas, esencialmente de acuerdo con PCR Protocols, A Guide to Methods and Applications, editado por Michael et al., Academic Press, 1990, utilizando la colección de línea celular o ADNc cromosómico apropiado para obtener el fragmento de la presente invención.

20 Un experto en la técnica puede diseñar fácilmente dichas sondas en base a la secuencia divulgada en el presente documento usando métodos de alineación por ordenador y análisis de secuencias conocidos en la técnica (cf. Molecular Cloning. A Laboratory Manual, segunda edición, editado por Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, 1989).

25 Las sondas de hibridación de la presente invención se pueden marcar por técnicas de marcado estándar tales como radiomarcado, marcado enzimático, marcado fluorescente, marcado biotina-avidina, quimioluminiscencia y similares. Después de la hibridación, se pueden visualizar las sondas usando métodos conocidos.

30 Las sondas de ácido nucleico de la presente invención incluyen sondas de ARN, así como de ADN, generándose dichas sondas usando técnicas conocidas en la técnica.

En una realización del método descrito anterior, se inmoviliza una sonda de ácido nucleico sobre un soporte sólido. Los ejemplos de dichos soportes sólidos incluyen, sin limitación, plásticos tales como policarbonato, carbohidratos complejos tales como agarosa y sefarosa, y resinas acrílicas, tales como poliacrilamida y perlas de látex. Las técnicas para acoplar sondas de ácido nucleico a dichos soportes sólidos son bien conocidas en la técnica.

Las muestras de prueba para métodos de sondaje de la presente invención incluyen, por ejemplo, células o extractos de ácido nucleico de células, o fluidos biológicos. La muestra usada en los métodos descritos anteriormente variará en base al formato de ensayo, el método de detección y la naturaleza de los tejidos, células o extractos que se van a ensayar. Los métodos para preparar extractos de ácido nucleico de células son bien conocidos en la técnica y se pueden adaptar fácilmente para obtener una muestra que es compatible con el método utilizado.

IV. Un método de detección de la presencia de ácido nucleico de PCA3 en una muestra

En otra realización, la presente invención se refiere a un método de detección de la presencia de ácido nucleico de PCA3 en una muestra que comprende a) poner en contacto la muestra con la sonda de ácido nucleico descrita anteriormente, bajo condiciones de hibridación específicas de modo que se produzca la hibridación, y b) detectar la presencia de la sonda unida a la molécula de ácido nucleico. Un experto en la técnica podría seleccionar la sonda de ácido nucleico de acuerdo con técnicas conocidas en la técnica como se describe anteriormente. Las muestras que se van a probar incluyen pero no se deben limitar a muestras de ARN o ADN a partir de tejido humano.

V. Un kit para detectar la presencia de ácido nucleico de PCA3 en una muestra

También se divulga un kit para detectar la presencia de ácido nucleico de PCA3 en una muestra que comprende al menos un medio de recipiente que tiene dispuesto en él la sonda de ácido nucleico anteriormente descrita. En una realización preferida, el kit comprende además otros recipientes que comprenden uno o más de los siguientes: reactivos de lavado y reactivos que pueden detectar la presencia de sonda de ácido nucleico unida. Los ejemplos de reactivos de detección incluyen, pero sin limitación, sondas radiomarcadas, sondas marcadas enzimáticamente (peroxidasa de rábano picante, fosfatasa alcalina), and sondas marcadas con afinidad (biotina, avidina o esteptavidina).

En detalle, un kit compartimentado incluye cualquier kit en el que los reactivos estén contenidos en recipientes separados. Dichos recipientes incluyen pequeños recipientes de vidrio, recipientes de plástico o tiras de plástico o papel. Dichos recipientes permiten la transferencia eficaz de reactivos desde un compartimento hasta otro compartimento de modo que las muestras y los reactivos no se contaminen de forma cruzada y los agentes o soluciones de cada recipiente se puedan añadir de forma cuantitativa desde un compartimento hasta otro. Dichos recipientes incluirán un recipiente que aceptará la muestra de prueba, un recipiente que contiene la sonda o cebadores usados en el ensayo, recipientes que contienen los reactivos de lavado (tal como solución salina tamponada con fosfato, tampones Tris y similares), y recipientes que contienen los reactivos usados para detectar la sonda hibridada, anticuerpo unido, producto amplificado o similares.

Un experto en la técnica reconocerá fácilmente que las sondas de ácido nucleico descritas en la presente invención se pueden incorporar fácilmente en uno de los formatos de kit establecidos que son bien conocidos en la técnica.

VI. Construcciones de ADN que comprenden una molécula de ácido nucleico de PCA3 y células que contienen estas construcciones

También se divulga una molécula de ADN recombinante que comprende, de 5' a 3', un promotor eficaz para iniciar la transcripción en una célula huésped y las moléculas de ácido nucleico descritas anteriormente. También se divulga una molécula de ADN recombinante que comprende un vector y una molécula de ácido nucleico descrita anteriormente.

También se divulga una molécula de ácido nucleico que comprende una región de control transcripcional funcional en una célula, una secuencia complementaria con una secuencia de ARN que codifica una secuencia de aminoácidos correspondiente con el polipéptido descrito anteriormente, y una región de terminación transcripcional funcional en la célula.

Preferentemente, las moléculas descritas anteriormente son moléculas de ADN aisladas y/o purificadas.

También se divulga una célula u organismo no humano que contiene una molécula de ácido nucleico descrita anteriormente.

También se divulga que el péptido se purifica a partir de células que se han alterado para expresar el péptido.

Como se usa en el presente documento, se dice que una célula está "alterada para expresar un péptido deseado"

cuando se hace que la célula, a través de manipulación genética, produzca una proteína que normalmente no produce o que normalmente la célula produce a bajos niveles. Un experto en la técnica puede adaptar fácilmente procedimientos para introducir y expresar secuencias genómicas, de ADNc o bien sintéticas en células eucariotas o bien procariotas.

Se dice que una molécula de ácido nucleico, tal como ADN, "puede expresar" un polipéptido si contiene secuencias de nucleótidos que contienen información reguladora transcripcional y traduccional y dichas secuencias están "unidas de forma funcional" a secuencias de nucleótidos que codifican el polipéptido. Una unión funcional es una unión en la que las secuencias de ADN reguladoras y la secuencia de ADN que se intenta expresar están conectadas de forma que permiten la expresión de secuencia génica. La naturaleza precisa de las regiones reguladoras necesarias para la expresión de secuencia génica puede variar de un organismo a otro, pero, en general, incluirá una región promotora que, en procariotas, contiene tanto el promotor (que dirige la iniciación de la transcripción de ARN) así como las secuencias de ADN que, cuando se transcriben en ARN, señalarán la iniciación de la síntesis. Normalmente, dichas regiones incluirán las secuencias no codificantes en 5' implicadas con la iniciación de la transcripción y traducción, tal como la caja TATA, secuencia de casquete, secuencia CAAT y similares.

Si se desea, se puede obtener la región no codificante en 3' para la secuencia codificante de PCA3 por los métodos descritos anteriormente. Esta región se puede conservar por sus secuencias reguladoras de terminación transcripcional, tal como terminación y poliadenilación. Por tanto, conservando la región 3' contigua de forma natural con la secuencia de ADN que codifica un gen PCA3, se pueden proporcionar las señales de terminación transcripcional. Cuando las señales de terminación transcripcional no son funcionales de forma satisfactoria en la célula huésped de expresión, entonces se puede sustituir una región 3' funcional en la célula huésped.

Se dice que dos secuencias de ADN (tal como una secuencia de región promotora y una secuencia codificante de PCA3) se unen de forma funcional si la naturaleza de la unión entre las dos secuencias de ADN (1) no da como resultado la introducción de una mutación de desplazamiento de marco, (2) no interfiere con la capacidad de la secuencia de región promotora de dirigir la transcripción de una secuencia codificante de PCA3, o (3) no interfiere con la capacidad de la secuencia codificante de PCA3 que se va a transcribir por la secuencia de región promotora. Por tanto, una región promotora se uniría de forma funcional a una secuencia de ADN si el promotor pudiera efectuar la transcripción de esa secuencia de ADN.

La presente invención engloba la expresión de la secuencia codificante de PCA3 (o un derivado funcional de la misma) en células procariotas o bien eucariotas. Los huéspedes procariotas, en general, son los más eficaces y convenientes para la producción de proteínas recombinantes y, por lo tanto, se prefieren para la expresión de la secuencia codificante de PCA3.

Lo más frecuentemente, los procariotas se representan por varias cepas de *E. coli*. Sin embargo, también se pueden usar otras cepas microbianas, incluidas otras cepas bacterianas. En sistemas procariotas, se pueden usar vectores plásmidos que contienen sitios de replicación y secuencias de control derivadas de una especie compatible con el huésped. Los ejemplos de vectores de plásmidos adecuados incluyen pBR322, pUC18, pUC19, pUC118, pUC119 y similares; vectores de fagos o bacteriófagos adecuados incluyen λ gt10, λ gt11 y similares; vectores víricos adecuados incluyen pMAM-neo, pKRC y similares. Preferentemente, el vector seleccionado de la invención tiene la capacidad para replicar en la célula huésped seleccionada.

Los huéspedes procariotas reconocidos incluyen bacterias tales como *E. coli*, *Bacillus*, *Streptomyces*, *Pseudomonas*, *Salmonella*, *Serratia* y similares. Sin embargo, bajo dichas condiciones, el péptido no se glucosilará. El huésped procariota debe ser compatible con el replicón y las secuencias de control en el plásmido de expresión.

Para expresar PCA3 en una célula procariota, es necesario unir de forma funcional la secuencia codificante de PCA3 a un promotor procariota funcional. Dichos promotores pueden ser constitutivos o bien, más preferentemente, regulables (es decir, inducibles o desreprimibles). Los ejemplos de promotores constitutivos incluyen el promotor *int* del bacteriófago λ , el promotor *bla* de la secuencia del gen β -lactamasa de pBR322, y el promotor CAT de la secuencia del gen cloramfenicol acetil transferasa de pBR325, y similares. Los ejemplos de promotores procariotas inducibles incluyen los promotores derecho e izquierdo principales del bacteriófago λ (P_L y P_R), el *trp*, los promotores *recA*, *lacZ*, *lacI*, y *gal* de *E. coli*, la α -amilasa (Ulmanen et al., J. Bacteriol. 162:176-182 (1985)) y los promotores específicos σ^{28} de *B. subtilis* (Gilman et al., Gene sequence 32:11-20 (1984)), los promotores de los bacteriófagos de *Bacillus* (Gryczan, en: The Molecular Biology of the Bacilli, Academic Press, Inc., NY (1982)), y promotores de *Streptomyces* (Ward et al., Mol. Gen. Genet. 203:468-478 (1986)). Los promotores procariotas se revisan por Glick (J. Ind. Microbiol. 1:277-282 (1987)); Cenatiempo (Biochimie 68:505-516 (1986)); y Gottesman (Ann. Rev. Genet. 18:415-442 (1984)).

La expresión apropiada en una célula procariota también requiere la presencia de un sitio de unión a ribosoma corriente arriba de la secuencia codificante de la secuencia génica. Dichos sitios de unión a ribosoma se divulgan, por ejemplo, por Gold et al. (Ann. Rev. Microbiol. 35:365-404 (1981)).

La selección de secuencias de control, vectores de expresión, métodos de transformación y similares, es independiente del tipo de célula huésped usada para expresar el gen. Como se usa en el presente documento, "célula"; "línea celular" y "cultivo celular" se pueden usar de forma intercambiable y dichas designaciones incluyen la progenie. Por tanto, las palabras "transformantes" o "células transformadas" incluyen la célula objeto principal y cultivos derivados de la misma, sin tener en cuenta el número de transferencias. También se entiende que toda la progenie puede no ser precisamente idéntica en el contenido de ADN, debido a mutaciones deliberadas o inadvertidas. Sin embargo, tal como se define, la progenie mutante tiene la misma funcionalidad que la de la célula originalmente transformada.

Las células huésped que se pueden usar en los sistemas de expresión de la presente invención no están estrictamente limitadas, siempre que sean adecuadas para su uso en la expresión del péptido de PCA3 de interés. Los huéspedes adecuados incluyen células eucariotas.

Los huéspedes eucariotas preferidos incluyen, por ejemplo, levaduras, hongos, células de insecto, células de mamífero *in vivo* o bien en cultivo tisular. Las células de mamífero preferidas incluyen células HeLa, células de origen fibroblástico tales como VERO o CHO-K1, o células de origen linfóide y sus derivados.

Además, también están disponibles células vegetales como huéspedes, y están disponibles secuencias de control compatibles con células vegetales, tales como las secuencias del virus del mosaico de la coliflor 35S y 19S, y del promotor de nopalina sintasa y de señal de poliadenilación.

Otro huésped preferido es una célula de insecto, por ejemplo *Drosophila* larvae. Usando células de insecto como huéspedes, se puede usar el promotor de alcohol deshidrogenasa *Drosophila*, Rubin, Science 240:1453-1459 (1988). De forma alternativa, se pueden diseñar vectores de baculovirus para expresar grandes cantidades de PCA3 en células de insectos (Jasny, Science 238:1653 (1987); Miller et al., In: Genetic Engineering (1986), Setlow, J.K., et al., eds., Plenum, vol. 8, p. 277-297).

Las diferentes células huésped tienen mecanismos característicos y específicos para el procesamiento traduccional y postraduccional y la modificación (por ejemplo, glucosilación, escisión) de proteínas. Se pueden elegir líneas celulares o sistemas huésped apropiados para garantizar la modificación y el procesamiento deseados de la proteína exógena expresada.

Se puede utilizar cualquiera de una serie de sistemas de expresión de secuencias génicas de levaduras que incorporan elementos promotores y de terminación de las secuencias génicas expresadas de forma activa que codifican enzimas glucolíticas. Estas enzimas se producen en grandes cantidades cuando las levaduras se hacen crecer en medios ricos en glucosa. Las secuencias génicas glucolíticas conocidas también pueden proporcionar señales de control transcripcional muy eficaces.

Las levaduras proporcionan ventajas sustanciales porque también pueden llevar a cabo modificaciones peptídicas post-traduccionales. Existen una serie de estrategias de ADN recombinante que utilizan secuencias promotoras fuertes y un número de copias alto de plásmidos que se pueden utilizar para la producción de las proteínas deseadas en levaduras. Las levaduras reconocen secuencias líder en productos de secuencias génicas de mamífero clonadas y segregan péptidos que llevan secuencias líder (es decir, pre-péptidos). Para un huésped de mamífero, están disponibles varios sistemas de vectores posibles para la expresión de PCA3.

Se puede emplear una amplia variedad de secuencias reguladoras transcripcionales y traduccionales, dependiendo de la naturaleza del huésped. Las señales reguladoras transcripcionales y traduccionales se pueden derivar de fuentes víricas, tales como adenovirus, virus de papiloma bovino, virus de simio o similares, donde las señales reguladoras se asocian con una secuencia génica particular que tiene un nivel alto de expresión. De forma alternativa, se pueden emplear promotores de productos de expresión de mamífero, tales como actina, colágeno, miosina, y similares. Se pueden seleccionar señales reguladoras de iniciación transcripcional para la expresión o activación, de modo que se pueda modular la expresión de las secuencias génicas. Son de interés señales reguladoras que son sensibles a la temperatura de modo que variando la temperatura, la expresión se puede reprimir o iniciar, o están sometidas a regulación química (tal como de metabolito).

Como se analiza anteriormente, la expresión de PCA3 en huéspedes eucariotas requiere el uso de regiones reguladoras eucariotas. En general, dichas regiones incluirán una región promotora suficiente para dirigir la iniciación de la síntesis de ARN. Los promotores eucariotas preferidos incluyen, por ejemplo, el promotor de la secuencia del gen metalotioneína I de ratón (Hamer et al., J. Mol. Appl. Gen. 1:273-288 (1982)); el promotor TK del virus de Herpes (McKnight, Cell 31:355-365 (1982)); el promotor temprano del SV40 (Benoist et al., Nature (Londres) 290:304-310 (1981)); la secuencia promotora del gen *gal4* de levadura (Johnston et al., Proc. Natl. Acad. Sci. (USA) 79:6971-6975 (1982); Silver et al., Proc. Natl. Acad. Sci. (USA) 81:5951-5955 (1984)) el promotor del gen inmediatamente temprano del VMC (Thomsen et al., Proc. Natl. Acad. Sci. (USA) 81:659-663 (1984)).

Como se conoce ampliamente, la traducción de ARNm eucariota se inicia en el codón que codifica la primera

metionina. Por este motivo, es preferible garantizar que la unión entre un promotor eucariota y una secuencia codificante de PCA3 no contiene ningún codón de intervención que pueda codificar una metionina (es decir, AUG). La presencia de dichos codones da como resultado la formación de una proteína de fusión (si el codón AUG está en el mismo marco de lectura que la secuencia codificante de PCA3) o una mutación de desplazamiento de marco (si el codón AUG no está en el mismo marco de lectura que la secuencia codificante de PCA3).

Se pueden introducir una molécula de ácido nucleico de PCA3 y un promotor unido de forma funcional en una célula procariota o eucariota receptora como una molécula de ADN (o ARN) que no se replica, que puede ser una molécula lineal o bien, más preferentemente, una molécula circular covalente cerrada. Puesto que dichas moléculas no pueden realizar una replicación autónoma, se puede producir la expresión del gen a través de la expresión transitoria de la secuencia introducida. De forma alternativa, se puede producir la expresión permanente a través de la integración de la secuencia de ADN introducida en el cromosoma huésped.

En una realización, se emplea un vector que puede integrar las secuencias génicas deseadas en el cromosoma de la célula huésped. Se pueden seleccionar células que tengan integrado de forma estable el ADN introducido en sus cromosomas introduciendo además uno o más marcadores que permitan la selección de células huésped que contengan el vector de expresión. El marcador puede proporcionar prototrofia a un huésped auxotrófico, resistencia a biocidas, por ejemplo, antibióticos, o metales pesados, tales como cobre, o similares. La secuencia génica del marcador seleccionable se puede unir directamente a las secuencias génicas de ADN para expresarse o bien introducirse en la misma célula por co-transfección. También son necesarios elementos adicionales para la síntesis óptima de ARNm de proteína de unión monocatenario. Estos elementos pueden incluir señales de ajuste, así como promotores de transcripción, secuencias señal potenciadoras y señales de terminación. Los vectores de expresión de ADNc que incorporan dichos elementos incluyen los descritos por Okayama, *Molec. Cell. Biol.* 3:280 (1983).

En una realización preferida, molécula de ácido nucleico introducida se incorporará en un plásmido o vector vírico que puede realizar replicación autónoma en el huésped receptor. Con este fin se puede emplear cualquiera de una amplia variedad de vectores. Los factores de importancia en la selección de un plásmido o vector vírico particular incluyen: la facilidad con la que las células receptoras que contienen el vector se pueden reconocer y seleccionar a partir de esas células receptoras que no contienen el vector; el número de copias del vector que se desean en un huésped particular; y si es deseable poder "lanzar" el vector entre células huésped de diferentes especies. Los vectores procariotas preferidos incluyen plásmidos tales como los que pueden realizar replicación en *E. coli* (tales como, por ejemplo, pBR322, ColE1, pSC101, pACYC 184, Π IX. Dichos plásmidos se divulgan, por ejemplo, por Sambrook (cf. *Molecular Cloning: A Laboratory Manual*, segunda edición, editado por Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, 1989). Los plásmidos de *Bacillus* incluyen pC194, pC221, pT127, y similares. Dichos plásmidos se divulgan por Gryczan (en: *The Molecular Biology of the Bacilli*, Academic Press, NY (1982), p. 307-329). Los plásmidos de *Streptomyces* adecuados incluyen pIJ101 (Kendall et al., *J. Bacteriol.* 169:4177-4183 (1987)), y bacteriófagos de *Streptomyces* tales como ϕ C31 (Chater et al., In: *Sixth International Symposium on Actinomycetales Biology*, Akademiai Kiado, Budapest, Hungría (1986), p. 45-54). Los plásmidos de *Pseudomonas* se revisan por John et al. (*Rev. Infect. Dis.* 8:693-704 (1986)), y Izaki (*Jpn. J. Bacteriol.* 33:729-742 (1978)).

Los plásmidos eucariotas preferidos incluyen, por ejemplo, BPV, vaccinia, SV40, círculo de 2 micras, y similares, o sus derivados. Dichos plásmidos son bien conocidos en la técnica (Botstein et al., *Miami Wntr. Symp.* 19:265-274 (1982); Broach. En: *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, p. 445-470(1981); Broach, *Cell* 28:203-204(1982); Bollon et al., *J. Clin. Hematol. Oncol.* 10:39-48 (1980); Maniatis, en: *Cell Biology: A Comprehensive Treatise*, vol. 3, Gene Sequence Expression, Academic Press, NY, p. 563-608 (1980)).

Una vez que se ha preparado el vector o la molécula de ácido nucleico que contiene la(s) construcción/construcciones para la expresión, se puede(n) introducir la(s) construcción/construcciones de ADN en una célula huésped apropiada por cualquiera de una variedad de medios adecuados, es decir, transformación, transfección, conjugación, fusión de protoplastos, electroporación, tecnología de pistola de partículas, precipitación de fosfato de calcio, microinyección directa, y similares. Después de la introducción del vector, se hacen crecer células receptoras en un medio selectivo, que selecciona el crecimiento de células que contienen el vector. La expresión de la(s) molécula(s) de gen clonado da como resultado la producción de PCA3. Esto puede tener lugar en las células transformadas como tal o después de la inducción de estas células para la diferenciación (por ejemplo, por la administración de bromodesoxiuracilo a células de neuroblastoma o similares).

VII. Un anticuerpo que tiene afinidad de unión a un polipéptido de PCA3 y un hibridoma que contiene el anticuerpo

También se divulga un anticuerpo que tiene afinidad de unión específicamente para un polipéptido de PCA3 como se describe anteriormente o específicamente para un fragmento de unión a polipéptido de PCA3 del mismo. Un anticuerpo se une específicamente a un polipéptido de PCA3 o fragmento de unión del mismo si no se une a polipéptidos distintos de PCA3. Los que se unen selectivamente a PCA3 se elegirían para su uso en métodos que podrían incluir, pero sin limitación, el análisis de la expresión de PCA3 alterada en tejido que contiene PCA3.

Se pueden usar las proteínas PCA3 como se divulga en el presente documento en varios procedimientos y métodos,

tales como para la generación de anticuerpos, para su uso en identificar composiciones farmacéuticas y para estudiar la interacción ADN/proteína.

Se pueden usar los péptidos de PCA3 como se divulga en el presente documento para producir anticuerpos o
5 hibridomas. Un experto en la técnica reconocerá que si se desea un anticuerpo, se generaría un péptido de este tipo, como se describe en el presente documento y se usa como inmunógeno.

Los anticuerpos como se divulga en el presente documento incluyen anticuerpos monoclonales y policlonales, así
10 como fragmentos de estos anticuerpos. La descripción incluye además anticuerpos monocatenarios. Se pueden generar fragmentos de anticuerpos que contengan el idiotipo de la molécula por técnicas conocidas. Por ejemplo, dichos fragmentos incluyen, pero sin limitación: el fragmento F(ab')₂; los fragmentos Fab', fragmentos Fab y fragmentos Fv.

Son de especial interés como se divulga en el presente documento anticuerpos para PCA3 que se producen en
15 seres humanos, o que están "humanizados" (es decir, no inmunógeno en un ser humano) por tecnología recombinante o de otro tipo. Se pueden producir anticuerpos humanizados, por ejemplo, reemplazando una porción inmunógena de un anticuerpo con una porción correspondiente, pero no inmunógena (es decir, anticuerpos quiméricos) (Robinson, R.R. et al., publicación de patente internacional PCT/US86/02269; Akira, K. et al., solicitud de patente europea 184.187; Taniguchi, M., solicitud de patente europea 171.496; Morrison, S.L. et al., solicitud de
20 patente europea 173.494; Neuberger, M.S. et al., solicitud PCT WO 86/01533; Cabilly, S. et al., solicitud de patente europea 125,023; Better, M. et al., Science 240:1041-1043 (1988); Liu, A.Y. et al., Proc. Natl. Acad. Sci. USA 84:3439-3443 (1987); Liu, A.Y. et al., J. Immunol. 139:3521-3526 (1987); Sun, L.K. et al., Proc. Natl. Acad. Sci. USA 84:214-218 (1987); Nishimura, Y. et al., Canc. Res. 47:999-1005 (1987); Wood, C.R. et al., Nature 314:446-4.49 (1985)); Shaw et al., J. Natl. cancer Inst. 80:1553-1559 (1988). Se proporcionan revisiones generales de
25 anticuerpos "humanizados" por Morrison, S.L. (Science, 229:1202-1207 (1985)) y por Oi, V.T. et al., BioTechniques 4:214 (1986)). Los anticuerpos "humanizados" adecuados se pueden producir de forma alternativa por sustitución de CDR o CEA (Jones, P.T. et al., Nature 321:552-525 (1986); Verhoeyan et al., Science 239:1534 (1988); Beidler, C.B. et al., J. Immunol. 141:4053-4060(1988)).

También se divulga un hibridoma que produce el anticuerpo monoclonal descrito anteriormente. Un hibridoma es
30 una línea celular inmortalizada que puede segregar un anticuerpo monoclonal específico.

En general, las técnicas para preparar anticuerpos monoclonales e hibridomas son bien conocidas en la técnica
35 (Campbell, "Monoclonal Antibody Technology. Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands (1984); St. Groth et al., J. Immunol. Methods 35:1-21 (1980)).

Se puede inmunizar cualquier animal (ratón, conejo y similares) que se sabe que produce anticuerpos con el
40 polipéptido seleccionado. Los métodos para la inmunización son bien conocidos en la técnica. Dichos métodos incluyen inyección subcutánea o interperitoneal del polipéptido. Un experto en la técnica reconocerá que la cantidad de polipéptido usado para la inmunización variará en base al animal que se inmuniza, la antigenicidad del polipéptido y el sitio de inyección.

El polipéptido se puede modificar o administrar en un coadyuvante para incrementar la antigenicidad del péptido. Los
45 métodos de incremento de la antigenicidad de un polipéptido son bien conocidos en la técnica. Tales procedimientos incluyen acoplar el antígeno con una proteína heteróloga (tal como globulina o β-galactosidasa) o a través de la inclusión de un coadyuvante durante la inmunización.

Para anticuerpos monoclonales, se retiran células del bazo de los animales inmunizados, se fusionan con células de
50 mieloma y se deja que se vuelvan células de hibridoma que producen anticuerpos monoclonales.

Se puede usar uno cualquiera de una serie de métodos bien conocidos en la técnica para identificar la célula de
55 hibridoma que produce un anticuerpo con las características deseadas. Estos incluyen cribar los hibridomas con un ensayo ELISA, análisis por transferencia de Western o radioinmunoensayo (Lutz et al., Exp. Cell Res. 175:109-124 (1988)).

Los hibridomas que segregar los anticuerpos deseados se clonan y se determina la clase y subclase usando
60 procedimientos conocidos en la técnica (Campbell, Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology, supra (1984)).

Para anticuerpos policlonales, se aíslan antiseros que contienen anticuerpos del animal inmunizado y se criba para
determinar la presencia de anticuerpos con la especificidad deseada usando uno de los procedimientos descritos anteriormente.

También se divulga que los anticuerpos descritos anteriormente se marcan de forma detectable. Los anticuerpos se
65 pueden marcar de forma detectable a través del uso de radioisótopos, marcas de afinidad (tales como biotina,

avidina y similares), marcas enzimáticas (tales como peroxidasa de rábano picante, fosfatasa alcalina y similares) marcas fluorescentes (tales como FITC o rodamina, y similares), átomos paramagnéticos, y similares. Procedimientos para llevar a cabo dicho marcado son bien conocidos en la técnica, por ejemplo, véanse (Sternberger et al., J. Histochem. Cytochem. 18:315 (1970); Bayer et al., Meth. Enzym. 62:308 (1979); Engval et al., Immunol. 109:129 (1972); Goding, J. Immunol. Meth. 13:215 (1976)). Se pueden usar los anticuerpos marcados como se divulga en el presente documento para ensayos *in vitro*, *in vivo* e *in situ* para identificar células o tejidos que expresan un péptido específico.

En otra realización como se divulga en el presente documento, los anticuerpos descritos anteriormente se inmovilizan sobre un soporte sólido. Los ejemplos de dichos soportes sólidos incluyen plásticos tales como policarbonato, carbohidratos complejos tales como agarosa y sefarosa, resinas acrílicas y tal como poliacrilamida y perlas de látex. Las técnicas para acoplar anticuerpos a dichos soportes son bien conocidas en la técnica (Weir et al "Handbook of Experimental Immunology" 4ª Ed., Blackwell Scientific Publications, Oxford, Inglaterra, capítulo 10 (1986); Jacoby et al., Meth. Enzym. 34 Academic Press, N.Y. (1974)). Se pueden usar los anticuerpos inmovilizados de la presente invención para ensayos *in vitro*, *in vivo* e *in situ* así como en inmunocromatografía.

Además, un experto en la técnica puede adaptar fácilmente procedimientos actualmente disponibles, así como las técnicas, métodos y kits divulgados anteriormente con respecto a anticuerpos, para generar péptidos que se pueden unir a una secuencia peptídica específica para generar péptidos antipéptido diseñados de forma racional, por ejemplo, por ejemplo, véase Hurby et al., "Application of Synthetic Peptides: Antisense Peptides", en Synthetic Peptides, A User's Guide, W.H. Freeman, NY, p. 289-307 (1992), y Kaspiczak et al., Biochemistry 28:9230-8 (1989).

Se pueden generar péptidos antipéptido en una de dos formas. En primer lugar, se pueden generar péptidos antipéptido reemplazando los residuos de aminoácidos básicos que se encuentran en la secuencia peptídica de PCA3 con residuos ácidos, mientras que mantienen grupos polares hidrófobos y no cargados. Por ejemplo, los residuos de lisina, arginina y/o histidina se reemplazan con ácido aspártico o ácido glutámico y los residuos de ácido glutámico se reemplazan por lisina, arginina o histidina.

VIII. Un método de detección de un polipéptido o anticuerpo de PCA3 en una muestra

También se divulga un método de detección de un polipéptido de PCA3 en una muestra, que comprende: a) poner en contacto la muestra con un anticuerpo descrito anteriormente (o proteína), bajo condiciones tales que se formen inmunocomplejos, y b) detectar la presencia del anticuerpo unido al polipéptido. En detalle, los métodos comprenden incubar una muestra de prueba con uno o más de los anticuerpos como se divulga en el presente documento y probar si el anticuerpo se une a la muestra de prueba. Los niveles alterados de PCA3 en una muestra en comparación con niveles normales pueden indicar una enfermedad específica (por ejemplo, cáncer de próstata).

También se divulga un método de detección de un anticuerpo de PCA3 en una muestra que comprende: a) poner en contacto la muestra con una proteína PCA3 descrita anteriormente, bajo condiciones tales que se formen inmunocomplejos, y b) detectar la presencia de la proteína unida al anticuerpo o anticuerpo unido a la proteína. En detalle, los métodos comprenden incubar una muestra de prueba con una o más de las proteínas de la presente invención y someter a ensayo si el anticuerpo se une a la muestra de prueba.

Las condiciones para incubar un anticuerpo con una muestra de prueba varían. Las condiciones de incubación dependen del formato empleado en el ensayo, los métodos de detección empleados y el tipo y la naturaleza del anticuerpo usado en el ensayo. Un experto en la técnica reconocerá que uno cualquiera de los formatos de ensayo inmunológicos comúnmente disponibles (tales como radioinmunoensayos, ensayos inmunoabsorbentes ligados a enzimas, prueba de Ouchterlony basada en difusión o ensayos inmunofluorescentes rápidos) se puede adaptar fácilmente para emplear los anticuerpos como se divulga en el presente documento. Ejemplos de dichos ensayos se pueden encontrar en Chard, An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Ámsterdam, Países Bajos (1986); Bullock et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Ámsterdam, Países Bajos (1985).

Las muestras de prueba de ensayos inmunológicos como se divulgan en el presente documento incluyen células, extractor de proteínas o membranas de células o fluidos biológicos tales como sangre, suero, plasma u orina. La muestra de prueba usada en el método descrito anteriormente variará en base al formato de ensayo, la naturaleza del método de detección y los tejidos, células o extractos usados como muestra que se va a someter a ensayo. Los métodos para preparar extractos de proteínas o extractos de membranas son bien conocidos en la técnica y se pueden adaptar fácilmente para obtener una muestra que se pueda usar con el sistema utilizado.

IX. Un kit de diagnóstico que comprende proteína o anticuerpo de PCA3

También se divulga un kit que contiene todos los reactivos necesarios para llevar a cabo los métodos de detección descritos previamente.

El kit puede comprender: i) un primer medio de recipiente que contiene un anticuerpo descrito anteriormente, y ii) un segundo medio de recipiente que contiene un conjugado que comprende un compañero de unión del anticuerpo y una marca.

El kit puede comprender: i) un primer medio de recipiente que contiene una proteína descrita anteriormente, y preferentemente, ii) un segundo medio de recipiente que contiene un conjugado que comprende un compañero de unión de la proteína y una marca. Más específicamente, un kit de diagnóstico comprende proteína PCA3 como se describe anteriormente, para detectar anticuerpos en el suero de animales o seres humanos potencialmente infectados.

También se divulga que el kit comprende además uno o más recipientes que comprenden uno o más de los siguientes: reactivos de lavado y reactivos que pueden detectar la presencia de anticuerpos unidos. Los ejemplos de reactivos de detección incluyen, pero sin limitación, anticuerpos secundarios marcados, o como alternativa, si el anticuerpo primario está marcado, los reactivos de unión a anticuerpo, enzimáticos o cromofóricos que pueden reaccionar con el anticuerpo marcado. El kit compartimentado puede ser como se describe anteriormente para kits de sondas de ácido nucleico.

Un experto en la técnica reconocerá fácilmente que los anticuerpos descritos en el presente documento se pueden incorporar fácilmente en uno de los formatos de kit establecidos que son bien conocidos en la técnica.

X. Cribado de diagnóstico

Se debe entender que aunque el siguiente análisis se refiere específicamente a pacientes humanos, las enseñanzas también son aplicables a cualquier animal que exprese PCA3.

Los métodos de diagnóstico y cribado de la invención son especialmente útiles para un paciente que se sospecha que está en riesgo de desarrollar una enfermedad asociada con un nivel de expresión alterado de PCA3 en base a sus antecedentes familiares, o un paciente en el que se desea diagnosticar una enfermedad relacionado con PCA3 (por ejemplo, cáncer de próstata).

De acuerdo con la invención, ahora es posible el cribado presintomático de un individuo que necesita dicho cribado usando ADN que codifica la proteína PCA3 o el gen PCA3 como se divulga en el presente documento o fragmentos del mismo. El método de cribado de la invención permite el diagnóstico presintomático, incluido el diagnóstico prenatal, de la presencia de un gen PCA3 que falta o aberrante en individuos, y por tanto una opinión acerca de la probabilidad de que dicho individuo podría desarrollar o ha desarrollado una enfermedad asociada a PCA3. Esto es especialmente valioso para la identificación de vehículos de genes PCA3 que faltan o alterados, por ejemplo, de individuos con antecedentes familiares de una enfermedad asociada a PCA3. También se desea el diagnóstico temprano para maximizar una intervención oportuna apropiada.

En una realización preferida del método de cribado, se podría tomar una muestra de tejido de dicho individuo, se cribaría para detectar (1) la presencia del gen PCA3 "normal"; (2) la presencia de ARNm de PCA3 y/o (3) la presencia de proteína PCA3. Se puede caracterizar el gen humano normal en base, por ejemplo, a la detección de patrones de digestión por restricción en ADN "normal" frente al del paciente, incluido el análisis de RFLR, usando sondas de ADN preparadas frente a la secuencia de PCA3 (o un fragmento funcional del mismo) enseñado en la invención. De forma similar, se puede caracterizar el ARNm de PCA3 y compararlo con (a) niveles y/o (b) el tamaño de ARNm de PCA3 normal como se encuentra en una población humana que no está en riesgo de desarrollar una enfermedad asociada a PCA3 usando sondas similares. Por último, la proteína PCA3 se puede (a) detectar y/o (b) cuantificar usando un ensayo biológico para determinar la actividad de PCA3 o usando un ensayo inmunológico y anticuerpos de PCA3. Cuando se somete a ensayo la proteína PCA3, se prefiere el ensayo inmunológico por su velocidad. Un (1) patrón de tamaño de ADN de PCA3 aberrante, y/o (2) tamaños o niveles de ARNm de PCA3 aberrantes y/o (3) niveles de proteína PCA3 aberrante podrían indicar que el paciente está en riesgo de desarrollar una enfermedad asociada a PCA3.

Más específicamente, en el presente documento se proporciona un método de diagnóstico de la presencia o predisposición a desarrollar cáncer de próstata en un paciente.

Los métodos de cribado y de diagnóstico de la invención no requieren que se use la secuencia codificante de ADN de PCA3 completa para la sonda. Más precisamente, sólo es necesario usar un fragmento o longitud de ácido nucleico que sea suficiente para detectar la presencia del gen PCA3 en una preparación de ADN de un individuo normal o afectado, la ausencia de dicho gen, o una propiedad física alterada de dicho gen (tal como un cambio en el patrón de migración electroforético). Preferentemente, se usa cualquiera de las sondas como se describe anteriormente.

Se puede realizar el diagnóstico prenatal cuando se desee, usando cualquier método conocido para obtener células fetales, incluidos amniocentesis, muestreo de vellosidades coriónicas (CVS) y fetoscopia. Se puede usar análisis de

cromosoma para determinar si la porción del cromosoma que posee el gen PCA3 normal está presente en un estado heterocigótico.

XI. Tratamientos terapéuticos

5

A. Ácidos nucleicos terapéuticos

Un ácido nucleico terapéutico como agente terapéutico puede tener, pero sin limitación, al menos uno de los siguientes efectos terapéuticos sobre una célula diana: inhibición de la transcripción de una secuencia de ADN; inhibición de la traducción de una secuencia de ARN; inhibición de la transcripción inversa de una secuencia de ARN o ADN; inhibición de una modificación post-traduccional de una proteína; inducción de la transcripción de una secuencia de ADN; inducción de la traducción de una secuencia de ARN; inducción de la transcripción inversa de una secuencia de ARN o ADN; inducción de una modificación post-traduccional de una proteína; transcripción del ácido nucleico como ARN; traducción del ácido nucleico como proteína o enzima; e incorporación del ácido nucleico en un cromosoma de una célula diana para la expresión constitutiva o transitoria del ácido nucleico terapéutico.

Los efectos terapéuticos de los ácidos nucleicos terapéuticos incluyen, pero sin limitación: desactivar un gen defectuoso o procesar la expresión del mismo, tal como ARN o ADN antisentido; inhibir la replicación vírica o síntesis; tratamiento génico como expresión de un ácido nucleico heterólogo que codifica una proteína terapéutica o corregir una proteína defectuosa; modificar un defecto o subexpresión de un ARN tal como un ARN^{nh}, un ARN^m, un ARN^t, o un ARN^r; que codifica un fármaco o profármaco, o una enzima que genera un compuesto como fármaco o profármaco en células patológicas o normales que expresan el receptor quimérico; y cualquier otro efecto terapéutico conocido.

En el método de tratamiento de una enfermedad asociada a PCA3 (preferentemente, cáncer de próstata) en un paciente que necesita dicho tratamiento, se puede proporcionar un gen PCA3 que no es indicativo de un estado de enfermedad a las células de dicho paciente de una manera y en una cantidad que permita la expresión de la proteína PCA3 proporcionada por dicho gen, durante un tiempo y en una cantidad suficiente para tratar dicho paciente. Preferentemente, se usa tecnología de reemplazo génico ("desactivación") que podría reemplazar el gen PCA3 que provoca la enfermedad por un gen PCA3 que no provoca enfermedad (específicamente, cáncer de próstata).

Se incluyen también en la divulgación composiciones farmacéuticas que comprenden una cantidad eficaz de al menos un oligonucleótido antisentido de PCA3, en combinación con un vehículo farmacéuticamente aceptable. Dichos oligonucleótidos antisentido incluyen, pero sin limitación, al menos una secuencia de nucleótidos de 12-500 bases de longitud que es complementaria con los exones de PCA3 1, 2, 3, 4a-4d; una secuencia de ADN de SEQ ID NO:1, 3, 4, o 6; o una secuencia de ADN que codifica al menos 4 aminoácidos de SEQ ID NO:2 o SEQ ID NO:7.

De forma alternativa, se puede combinar el ácido nucleico de PCA3 con un vehículo lipófilo tal como uno cualquiera de un número de esteroides incluidos colesterol, colato y ácido desoxicólico. Un esteroide preferido es colesterol.

Se pueden administrar los ácidos nucleicos y las composiciones farmacéuticas de tratamiento génico con PCA3 como se divulga en el presente documento por cualquier medio que logre su fin destinado. Por ejemplo, la administración puede ser por vía parenteral, subcutánea, intravenosa, intramuscular, intraperitoneal o transdérmica. La dosificación administrada dependerá de la edad, salud y peso del receptor, tipo de tratamiento concurrente, si lo hay, frecuencia de tratamiento, y la naturaleza del efecto deseado.

Las composiciones como se divulgan en el presente documento incluyen todas las composiciones en las que está contenido el oligonucleótido antisentido de PCA3 en una cantidad eficaz para lograr una disminución en la expresión de al menos un gen PCA3. Aunque las necesidades individuales varían, la determinación de los intervalos óptimos de cada componente está dentro de la técnica. Normalmente, el ácido nucleico de PCA3 se puede administrar a mamíferos, por ejemplo seres humanos, en una dosis de 0,005 a 1 mg/kg/día, o una cantidad equivalente de la sal farmacéuticamente aceptable del mismo, por día de peso corporal del mamífero que se está tratando.

Las formulaciones adecuadas para la administración parenteral incluyen soluciones acuosas del ácido nucleico de PCA3 en forma soluble en agua, por ejemplo, sales solubles en agua. Además, se pueden administrar suspensiones de los compuestos activos como suspensiones de inyección oleosa apropiadas. Los disolventes o vehículos lipófilos adecuados incluyen aceites grasos, por ejemplo, aceite de sésamo o ésteres de ácidos grasos sintéticos, por ejemplo, oleato de etilo o triglicéridos. Las suspensiones de inyección acuosas pueden contener sustancias que incrementan la viscosidad de la suspensión, incluyen, por ejemplo, carboximetil celulosa de sodio, sorbitol y/o dextrano. Opcionalmente, la suspensión también puede contener estabilizantes.

Se conocen muchos sistemas de vectores en la técnica para proporcionar dicha administración a pacientes humanos que necesitan un gen o proteína que falta de la célula. Por ejemplo, se pueden usar sistemas de retrovirus, en especial sistemas de retrovirus modificados y en especial sistemas de virus del herpes simple (Gage et al., patente de los Estados Unidos N.º 5.082.670). Dichos métodos se proporcionan, por ejemplo, en las enseñanzas de

Breakfield, X.A. et al., *The New Biologist* 3:203-218 (1991); Huang, Q. et al., *Experimental Neurology* 115:303-316 (1992); documento WO93/03743; documento WO90/0944; Taylor, documento WO 92/06693; Mulligan, R.C., *Science* 260:926-932 (1993); y Brown et al., "Retroviral Vectors", en *DNA Cloning: A Practical Approach*, volumen 3, IRL Press, Washington, D.C. (1987). La administración de una secuencia de ADN que codifica una proteína PCA3 expresada normalmente reemplazará eficazmente el gen PCA3 responsable del estado de enfermedad (por ejemplo, cáncer de próstata).

Los medios por los que se puede introducir el vector que lleva el ácido nucleico dentro de la célula incluyen, pero sin limitación, microinyección, electroporación, transducción o transfección usando DEAE-dextrano, lipofección, fosfato de calcio u otros procedimientos conocidos por un experto en la técnica (*Molecular Cloning, A Laboratory Manual*, Sambrook et al., eds., Cold Spring Harbor Press, Plainview, New York (1989)).

También se divulga que un gen PCA3 normal se expresa como un gen recombinante en una célula, de modo que se pueden trasplantar las células en un mamífero, preferentemente un ser humano que necesite el tratamiento génico. Para proporcionar el tratamiento génico a un individuo, una secuencia genética que codifica todo o parte del gen PCA3 se inserta en un vector y se introduce en una célula huésped.

Otros métodos de tratamiento génico que se pueden usar para transferir ácido nucleico a un paciente se establecen en Chatterjee and Wong, *Current Topics in Microbiol. and Immuno.*, 218: 61-73 (1996); Zhang, J. *Mol. Med.* 74:191-204 (1996); Schmidt-Wolf y Schmidt-Wolf, *J. of Hematotherapy* 4:551-561 (1995); Shaughnessy et al., *Seminars in Oncology* 23(1): 159-171 (1996); y Dunbar *Annu. Rev. Med.* 47:11-20 (1996).

Se puede conferir especificidad para la expresión génica en células de cáncer de próstata usando secuencias reguladoras específicas de célula apropiadas, tales como potenciadores y promotores específicos de célula.

Por tanto, se puede usar el tratamiento génico para aliviar la patología relacionada con PCA3 inhibiendo la expresión inapropiada de una forma particular de PCA3. Además, se puede usar el tratamiento génico para aliviar dichas patologías proporcionando el nivel de expresión apropiada de una forma particular de PCA3. En este caso, las secuencias de ácido nucleico de PCA3 particulares se pueden codificar por construcciones de ADN o ARN que se administran en forma de virus, como se describe anteriormente.

B. Antagonistas y agonistas de PCA3

Se puede evaluar la capacidad de antagonistas y agonistas de PCA3 para interferir o potenciar la actividad de PCA3 con células que contienen PCA3. Se puede usar un ensayo para determinar la actividad de PCA3 en células para determinar la funcionalidad de la proteína PCA3 en presencia de un agente que puede actuar como antagonista o agonista, y por tanto, se identifican agentes que interfieren o potencian la actividad de PCA3.

Los agentes cribados en los ensayos pueden ser, pero sin limitación, anticuerpos, péptidos, carbohidratos, derivados de vitaminas u otros agentes farmacéuticos. Estos agentes se pueden seleccionar y cribar 1) al azar, 2) por una selección racional o 3) por diseño usando, por ejemplo, técnicas de modelado de proteínas o ligandos (preferentemente, modelado por ordenador).

Para el cribado al azar, los agentes tales como anticuerpos, péptidos, carbohidratos, agentes farmacéuticos y similares se seleccionan al azar y se someten a ensayo para determinar su capacidad de unión a o estimular/bloquear la actividad de la proteína PCA3.

De forma alternativa, los agentes se pueden seleccionar o diseñar racionalmente. Como se usa en el presente documento, se dice que un agente se "selecciona o diseña racionalmente" cuando se elige el agente en base a la configuración de la proteína PCA3.

También se divulga un método de cribado de un antagonista o agonista que estimule o bloquee la actividad de PCA3 que comprende:

(a) incubar una célula que expresa PCA3 con un agente que se va a someter a prueba; y

(b) someter a ensayo la célula para determinar la actividad de la proteína PCA3 midiendo el efecto de los agentes sobre la unión a ATP de PCA3.

Se puede usar cualquier célula en el ensayo anterior siempre que exprese una forma funcional de PCA3 y se puede medir la actividad de PCA3. Las células de expresión preferidas son células u organismos eucariotas. Se pueden modificar dichas células para que contengan secuencias de ADN que codifiquen PCA3 usando procedimientos de rutina conocidos en la técnica. De forma alternativa, un experto en la técnica puede introducir ARNm que codifique la proteína PCA3 directamente en la célula.

Usando ligandos de PCA3 (ligandos que incluyen los antagonistas y agonistas como se describe anteriormente) se

divulga adicionalmente un método para modular la actividad de la proteína PCA3 en una célula. En general, se pueden formular ligandos (antagonistas y agonistas) que se han identificado por bloquear o estimular la actividad de PCA3 de modo que se puede poner en contacto el ligando con una célula que exprese una proteína PCA3 *in vivo*. El contacto de una célula de este tipo con un ligando de este tipo da como resultado la modulación *in vivo* de la actividad de las proteínas PCA3. Mientras no exista una barrera de formulación ni una barrera de toxicidad, los ligandos identificados en los ensayos descritos anteriormente serán eficaces para su uso *in vivo*.

También se divulga un método de administración de PCA3 o ligando de PCA3 (incluidos antagonistas y agonistas de PCA3) a un animal (preferentemente, un mamífero (específicamente, un ser humano)) en una cantidad suficiente para lograr un nivel alterado de PCA3 en el animal. La PCA3 o el ligando de PCA3 administrado podría lograr funciones asociadas a PCA3. Además, puesto que la PCA3 se expresa en células de cáncer prostático, se podría usar la administración de PCA3 o ligando de PCA3 para alterar los niveles de PCA3 en dichas células.

Un experto en la técnica apreciará que se pueden determinar fácilmente las cantidades que se administrarán para cualquier protocolo de tratamiento particular. La dosificación no debe ser tan grande como para provocar efectos secundarios adversos, tales como reacciones cruzadas no deseadas, reacciones anafilácticas y similares. En general, la dosificación variará con la edad, condición, sexo y extensión de la enfermedad en el paciente, contraindicaciones, si las hay, y otras de dichas variables, que se ajustarán por el médico individual. La dosificación puede variar de desde 0,001 mg/kg hasta 50 mg/kg de PCA3 o ligando de PCA3, en una o más administraciones diarias, durante uno o varios días. Se puede administrar la PCA3 o el ligando de PCA3 por vía parenteral por inyección o por perfusión gradual con el tiempo. Se puede administrar por vía intravenosa, intraperitoneal, intramuscular o subcutánea.

Las preparaciones para la administración parenteral incluyen emulsiones, suspensiones y soluciones acuosas o no acuosas estériles. Ejemplos de disolventes no acuosos son propilenglicol, polietilenglicol, aceites vegetales tales como aceite de oliva, y ésteres orgánicos inyectables, tales como oleato de etilo. Los vehículos acuosos incluyen agua, soluciones alcohólicas/acuosas, emulsiones o suspensiones, incluidos medios salinos y tamponados. Los vehículos parenterales incluyen solución de cloruro de sodio, dextrosa de Ringer y cloruro de sodio, Ringer con lactato o aceites fijos. Los vehículos intravenosos incluyen reponedores de fluidos y nutrientes, reponedores de electrolitos, tales como los basados en la dextrosa de Ringer, y similares. También pueden estar presentes conservantes y otros aditivos tales como, por ejemplo, antimicrobianos, antioxidantes, agentes quelantes, gases inertes y similares. Véase, en general, Remington's Pharmaceutical Science, 16^a ed., Mack Eds. (1980).

También se divulga una composición farmacéutica que comprende PCA3 o ligando de PCA3 en una cantidad suficiente para alterar la actividad asociada a PCA3, y un diluyente, vehículo o excipiente farmacéuticamente aceptable. Se pueden determinar fácilmente concentraciones y tamaños de unidades de dosificación apropiados por un experto en la técnica como se describe anteriormente (véase, por ejemplo, Remington's Pharmaceutical Sciences (16^a ed., Osol, A., Ed., Mack, Easton PA (1980) y el documento WO 91/19008).

C. Inmunoterapia

También se divulgan los anticuerpos de PCA3 descritos anteriormente (preferentemente, anticuerpos murinos de PCA3 y anticuerpos murinos-humanos de PCA3 quiméricos, y fragmentos y regiones de los mismos) que inhiben o neutralizan la actividad biológica de PCA3 *in vivo* y son específicos para PCA3. Se pueden usar estos anticuerpos con fines terapéuticos en sujetos que tienen patologías o afecciones asociadas a la presencia de expresión de PCA3 aberrante. Los anticuerpos y fragmentos, regiones y derivados de los mismos, como se divulga en el presente documento contienen preferentemente al menos una región que reconoce un epítipo de PCA3 que tiene actividad biológica de inhibición y/o neutralización *in vivo*.

El tratamiento comprende administrar por vía parenteral una dosis individual o múltiples dosis del anticuerpo, fragmento o derivado. Son preferidos para un uso farmacéutico humano anticuerpos, fragmentos y regiones, quiméricos y murinos que inhiben y/o neutralizan la PCA3 potentes de alta afinidad, de esta descripción.

Se pueden administrar anticuerpos monoclonales como se divulga en el presente documento por cualquier medio que permita que el agente activo alcance el sitio de acción del agente en el cuerpo de un mamífero. Debido a que las proteínas son susceptibles a ser digeridas cuando se administran por vía oral, de forma ordinaria se usará la administración parenteral, es decir, intravenosa, subcutánea, intramuscular, para optimizar la absorción.

Se pueden administrar anticuerpos monoclonales como se divulga en el presente documento como agentes terapéuticos individuales o bien en combinación con otros agentes terapéuticos. Se pueden administrar solos, pero, en general, se administran con un vehículo farmacéutico seleccionado en base a la vía de administración elegida y la práctica farmacéutica estándar.

Por supuesto, la dosificación administrada variará dependiendo de factores conocidos tales como las características farmacodinámicas del agente particular, y su modo y vía de administración; edad, salud y peso del receptor; naturaleza y extensión de los síntomas, tipo de tratamiento concurrente, frecuencia de tratamiento, y el efecto

deseado. Por lo general, una dosificación diaria de ingrediente activo puede ser aproximadamente de 0,1 a 100 miligramos por kilogramo de peso corporal. Generalmente, de 0,5 a 50, y preferentemente de 1 a 10 miligramos por kilogramo por día dado en dosis divididas de 1 a 6 veces al día o en forma de liberación sostenida es eficaz para obtener resultados deseados.

En general, las formas de dosificación (composición) adecuadas para una administración interna contienen de desde aproximadamente 1 miligramo hasta aproximadamente 500 miligramos de ingrediente activo por unidad. En estas composiciones farmacéuticas, el ingrediente activo estará presente normalmente en una cantidad de aproximadamente el 0,5-95 % en peso en base al peso total de la composición.

Para la administración parenteral, el anticuerpo se puede formular como una solución, suspensión, emulsión o polvo liofilizado en asociación con un vehículo parenteral farmacéuticamente aceptable. Ejemplos de dichos vehículos son agua, solución salina, solución de Ringer, solución de dextrosa y seroalbúmina humana al 5%. También se pueden usar liposomas y vehículos no acuosos tales como aceites fijos. El vehículo o polvo liofilizado puede contener aditivos que mantienen la isotonicidad (por ejemplo, cloruro de sodio, manitol) y estabilidad química (por ejemplo, tampones y conservantes). La formulación se esteriliza por técnicas usadas comúnmente.

Los vehículos farmacéuticos adecuados se describen en la edición más reciente de Remington's Pharmaceutical Sciences, A. Osol, un texto de referencia estándar en el campo.

Los anticuerpos, fragmentos y regiones murinos y quiméricos como se divulga en el presente documento, sus fragmentos y derivados se pueden usar de forma terapéutica como inmunoconjugados (véase para su revisión: Dillman, R.O., Ann. Int. Med. 111:592-603 (1989)). Se pueden acoplar a proteínas citotóxicas, incluidas, pero sin limitación, ricina-A, toxina *Pseudomonas* y toxina *Diphtheria*. Las toxinas conjugadas a anticuerpos u otros ligandos, son conocidas en la técnica (véase, por ejemplo, Olsnes, S. et al., Immunol. Today 10:291-295 (1989)). Normalmente, las toxinas de plantas y bacterias matan células alterando la maquinaria sintética proteínica.

Los anticuerpos como se disponen en el presente documento se pueden conjugar a tipos adicionales de restos terapéuticos, incluidos pero sin limitación, radionúclidos, agentes citotóxicos y fármacos. Ejemplos de radionúclidos que se pueden acoplar a anticuerpos y suministrarse *in vivo* a sitios de antígeno incluyen ^{212}Bi , ^{131}I , ^{186}Re y ^{90}Y , lista que no pretende ser exhaustiva. Los radionúclidos ejercen su efecto citotóxico irradiando localmente las células, lo que conduce a varias lesiones intracelulares, como es conocido en la técnica de radioterapia.

Los fármacos citotóxicos, que se pueden conjugar con anticuerpos y posteriormente usarse para tratamiento *in vivo* incluyen, pero sin limitación, daunorubicina, doxorubicina, metotrexato y mitomicina C. Los fármacos citotóxicos interfieren con procesos celulares críticos incluidos la síntesis de ARN, ADN y proteínas. Para una exposición más completa de estas clases de fármacos que son conocidos en la técnica, y sus mecanismos de acción, véase Goodman, A.G., et al., Goodman y Gilman's THE PHARMACOLOGICAL BASIS OF THERAPEUTICS, 7ª ed., Macmillan Publishing Co., 1985.

Los anticuerpos como se divulga en el presente documento se pueden utilizar ventajosamente en combinación con otros anticuerpos, fragmentos o regiones, monoclonales o murinos o quiméricos, o con linfocinas o factores de crecimiento hemopoético, etc., que sirven para incrementar el número o la actividad de las células efectoras que interaccionan con los anticuerpos.

XII. Animales no humanos con PCA3 transgénicos

Métodos de generación de animales no humanos transgénicos

Los animales no humanos como se divulga en el presente documento comprenden cualquier animal que tiene una interrupción o alteración transgénica del/de los gen(es) endógeno(s) (animales desactivados) y/o en el genoma en el que se ha introducido uno o más transgenes que dirigen la expresión de la PCA3 humana. También se prefiere la introducción de ácidos nucleicos de PCA3 antisentido.

Dichos animales no humanos incluyen vertebrados tales como roedores, primates no humanos, ovejas, perros, vacas, anfibios, reptiles, etc. Los animales no humanos preferidos se seleccionan a partir especies de animales mamíferas no humanas, lo más preferentemente, animales de la familia de los roedores incluidos ratas y ratones, lo más preferentemente ratones.

Los animales transgénicos como se divulga en el presente documento son animales en los que se ha introducido por medios no naturales (es decir, por manipulación humana), uno o más genes que no se producen de forma natural en el animal, por ejemplo, genes exógenos, genes endógenos diseñados genéticamente, etc. Los genes introducidos no naturales, conocidos como transgenes, pueden ser de la misma especie que el animal o de una especie diferente pero que no se encuentren de forma natural en el animal en la configuración y/o en el locus cromosómico conferido por el transgén. Los transgenes pueden comprender secuencias de ADN exógenas, es decir, secuencias que no se encuentran normalmente en el genoma del animal huésped. De forma alternativa o adicional, los transgenes pueden

comprender secuencias de ADN endógenas que son anormales porque se han reordenado o mutado *in vitro* para alterar el patrón de expresión *in vivo* normal del gen, o para alterar o eliminar la actividad biológica de un producto génico endógeno codificado por el gen. (Watson, J.D., et al., en Recombinant DNA, 2ª ed., W.H. Freeman & Co., New York (1992), páginas 255-272; Gordon, J.W., Intl. Rev. Cytol. 115:171-229 (1989); Jaenisch, R., Science 240:1468-1474 (1989); Rossant, J., Neuron 2:323-334 (1990)).

Los animales no humanos transgénicos como se divulga en el presente documento se producen introduciendo transgenes en la línea germinal del animal no humano. Se usan células diana embrionarias en varias fases de desarrollo para introducir los transgenes de la invención. Se usan diferentes métodos dependiendo de la fase de desarrollo de la(s) célula(s) embrionaria(s).

1. La microinyección de cigotos es el método preferido para incorporar transgenes en los genomas del animal en el transcurso de la práctica de la invención. Un cigoto, un óvulo fertilizado que no ha sufrido fusión pronucleica ni división celular posterior, es la célula diana preferida para la microinyección de secuencias de ADN transgénicas. El pronúcleo macho murino alcanza un tamaño de aproximadamente 20 micrómetros de diámetro, una característica que permite la inyección reproducible de 1-2 picolitros de una solución que contiene secuencias de ADN transgénicas. El uso de un cigoto para la introducción de transgenes tiene la ventaja de que, en la mayoría de los casos, las secuencias de ADN transgénicas inyectadas se incorporarán en el genoma del animal huésped antes de la primera división celular (Brinster, et al., Proc. Natl. Acad. Sci. (USA) 82:4438-4442 (1985)). Como consecuencia, todas las células de los animales transgénicos resultantes (animales fundadores) llevan de forma estable un transgén incorporado en un locus genético particular, denominado alelo transgénico. El alelo transgénico demuestra la herencia mendeliana: la mitad de la descendencia resultante del cruce de un animal transgénico con un animal no transgénico heredará el alelo transgénico, de acuerdo con las leyes de Mendel de selección aleatoria.

2. También se puede usar la integración vírica para introducir los transgenes de la invención en un animal. Los embriones en desarrollo se cultivan *in vitro* hasta la fase de desarrollo conocido como blastocisto. En este momento, los blastómeros se pueden infectar con retrovirus apropiados (Jaenisch, R., Proc. Natl. Sci. (USA) 73:1260-1264 (1976)). La infección de los blastómeros se potencia por retirada enzimática de la zona pelúcida (Hogan, et al., en Manipulating the Mouse Embryo, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1986)). Los transgenes se introducen por medio de vectores víricos que normalmente son defectuosos por replicación pero que aún son competentes para la integración de secuencias de ADN asociadas a virus, incluidas secuencias de ADN transgénicas ligadas a dichas secuencias víricas, en el genoma del animal huésped (Jahner, et al., Proc. Natl. Acad. Sci. (USA) 82:6927-6931 (1985); Van der Putten, et al., Proc. Natl. Acad. Sci. (USA) 82:6148-6152 (1985)). La transfección se obtiene de forma fácil y eficaz por cultivo de blastómeros sobre una monocapa de células que producen el vector vírico que contiene el transgén (Van der Putten, et al., Proc. Natl. Acad. Sci. (USA) 82:6148-6152 (1985); Stewart, et al., EMBO Journal 6:383-388 (1987)). De forma alternativa, la infección se puede realizar en una fase tardía, tal como blastocelo (Jahner, D., et al., Nature 298:623-628 (1982)). En cualquier acontecimiento, la mayoría de los animales fundadores producidos por integración vírica serán mosaicos para el alelo transgénico; esto es, el transgén se incorpora sólo en un subconjunto de todas las células que forman el animal fundador transgénico. Además, se pueden producir múltiples acontecimientos de integración vírica en un único animal fundador, lo que genera múltiples alelos transgénicos que se segregarán en futuras generaciones de descendencia. La introducción de transgenes en células de línea germinal por este método es posible pero probablemente se produzca a una frecuencia baja (Jahner, D., et al., Nature 298:623-628 (1982)). Sin embargo, una vez se ha introducido el transgén en las células de línea germinal por este método, se puede producir la descendencia en la que esté presente el alelo transgénico en todas las células del animal, es decir, en células tanto somáticas como de línea germinal.

3. Las células madre embrionarias (ES) también pueden servir como células diana para la introducción de los transgenes de la invención en animales. Las células ES se obtienen a partir de la preimplantación de embriones que se cultivan *in vitro* (Evans, M.J., et al., Nature 292:154-156 (1981); Bradley, M. O., et al., Nature 309:255-258 (1984); Gossler, et al., Proc. Natl. Acad. Sci. (USA) 83:9065-9069 (1986); Robertson et al., Nature 322:445-448 (1986); Robertson, E.J., en Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, E.J., ed., IRL Press, Oxford (1987), páginas 71-112). Las células ES, que están disponibles comercialmente (por ejemplo, de Genome Systems, Inc., St. Louis, MO), se pueden transformar con uno o más transgenes por métodos establecidos (Lovell-Badge, R.H., en Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, E.J., ed., IRL Press, Oxford (1987), páginas 153-182). Las células ES transformadas se pueden combinar con un blastocisto animal, donde después las células ES colonizan el embrión y contribuyen a la línea germinal del animal resultante, que es una quimera (compuesta de células derivadas de dos o más animales) (Jaenisch, R., Science 240:1468-1474 (1988); Bradley, A., en Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, E.J., ed., IRL Press, Oxford (1987), páginas 113-151). De nuevo, una vez se ha introducido el transgén en las células de línea germinal por este método, se puede producir la descendencia en la que esté presente el alelo transgénico en todas las células del animal, es decir, en células tanto somáticas como de línea germinal.

Como sea que se produzca, la introducción inicial de un transgén es un acontecimiento lamarckiano (no mendeliano). Sin embargo, los transgenes como se divulga en el presente documento se pueden integrar de forma estable en células de línea germinal y transmitirse a la descendencia del animal transgénico como locus mendelianos. Otras técnicas transgénicas dan como resultado animales transgénicos mosaico, en los que algunas

células llevan los transgenes y otras no. En animales transgénicos mosaico en los que las células de línea germinal no llevan los transgenes, no se produce la transmisión de los transgenes a la descendencia. No obstante, los animales transgénicos mosaico pueden demostrar fenotipos asociados con los transgenes.

- 5 Se pueden introducir transgenes en animales no humanos para proporcionar modelos animales para enfermedades humanas. Los transgenes que resultan de dichos modelos animales incluyen, por ejemplo, transgenes que codifican productos génicos mutantes asociados con un error congénito del metabolismo en una enfermedad genética humana y los transgenes que codifican un factor humano requerido para conferir susceptibilidad a un patógeno humano (es decir, una bacteria, virus u otro microorganismo patógeno) (Leder et al., patente de los EE. UU. 5.175.383 (29 de diciembre de 1992); Kindt et al., patente de los EE. UU. 5.183.949 (2 de febrero de 1993); Small et al., Cell 46:13-18 (1986); Hooper et al., Nature 326:292-295 (1987); Stacey et al., Nature 332:131-136 (1988); Windle et al., Nature 343:665-669 (1990); Katz et al., Cell 74:1089-1100 (1993)). Las mutaciones introducidas de forma transgénicas comprenden alelos nulos ("desactivados") en los que se sustituyen una secuencia de ADN que codifica un marcador seleccionable y/o detectable con una secuencia genética normalmente endógena para un animal no humano. Se pueden usar animales no humanos transgénicos resultantes que están predispuestos a una enfermedad, o en los que el transgén provoca una enfermedad, para identificar composiciones que inducen la enfermedad y para evaluar el potencial patógeno de composiciones conocidas o que se sospecha que inducen la enfermedad (Bems, A.J.M., patente de los EE. UU. 5.174.986 (29 de diciembre de 1992)), o para evaluar composiciones que se puede usar para tratar la enfermedad o mejorar los síntomas de la misma (Scott et al., documento WO 94/12627 (1994)).

La descendencia que ha heredado los transgenes como se divulga en el presente documento se distingue de la camada que no ha heredado los transgenes por análisis de material genético de la descendencia para determinar la presencia de biomoléculas que comprenden secuencias únicas correspondientes a secuencias de, o codificadas por, los transgenes como se divulga en el presente documento. Por ejemplo, los fluidos biológicos que contienen polipéptidos codificados únicamente por el marcador seleccionable de los transgenes como se divulga en el presente documento se puede someter a inmunoensayo para determinar la presencia de los polipéptidos. Un medio más sencillo y fiable de identificar una descendencia transgénica comprende obtener una muestra de tejido de una extremidad del animal, por ejemplo, una cola, y analizar la muestra para determinar la presencia de secuencias de ácido nucleico correspondientes a la secuencia de ADN de una única porción o porciones de los transgenes como se divulga en el presente documento, tal como el marcador seleccionable del mismo. Se puede determinar la presencia de dichas secuencias de ácido nucleico, por ejemplo, por análisis de hibridación ("Southern") con secuencias de ADN correspondientes a porciones únicas del transgén, análisis de los productos de reacciones de PCR usando secuencias de ADN en una muestra como sustratos y oligonucleótidos derivados de la secuencia de ADN del transgén, etc.

La presente invención se describe con más detalle en los siguientes ejemplos no limitantes.

Ejemplo 1

40 Aislamiento y caracterización de ADNc de PCA3 y ADN genómico

Para identificar nuevos marcadores para el cáncer de próstata, se usó un análisis de presentación diferencial (Liang et al., Science 257: 967-971 (1992)) para identificar genes sobreexpresados en carcinomas prostáticos en comparación con próstata normal; se extrajo ARN total de tejido prostático normal, benigno hiperplásico y maligno del mismo paciente. Usando veinte combinaciones diferentes de cebadores (cuatro cebadores anclados, cinco cebadores arbitrarios), se identificaron once ARNm aparentemente expresados de forma diferencial (es decir, sobreexpresados de forma consistente en todos los carcinomas estudiados y no expresados en tejido normal o BPH). Se usaron los fragmentos de ADN complementario (ADNc) como sondas para el análisis de bandas de Northern para confirmar la sobreexpresión consistente en los tumores prostáticos usados para la presentación diferencial. Una de las sondas (denominada DD3, un ADNc de 486 pb) detectó dos transcritos principales (de 2,3 y 4,0 kb) que se sobreexpresan altamente en 47 de 50 tumores prostáticos humanos estudiados, mientras que no se encontró expresión de estos transcritos en tejido normal o BPH de los mismos pacientes.

55 Para obtener un clon de ADNc de longitud completa, se construyó una colección de ADNc usando ARNm aislado de tejido de tumor prostático primario humano. Se obtuvieron 250 clones relacionados con DD3 positivos a partir del cribado de esta colección. Se purificaron 80 clones y se determinó la secuencia de nucleótidos de estos clones por análisis de secuencia automatizado.

60 Se cribó una colección genómica construida de ADN genómico de placenta humana clonado en λ FIX2 usando DD3 como sonda. Se obtuvieron cuatro clones diferentes, dos de ellos situados hacia el extremo 5' del gen (AFIX-ME3 y -ME4) y dos clones situados hacia el extremo 3' del gen (XFIX-ME1 y -ME2). Se subclonó el extremo 5' de XFIX-ME4 y se usó como sonda para cribar la colección genómica. Se aislaron tres clones nuevos, únicos (λ FIX-IH1, IH2 y IH6).

65 De los 80 clones de ADNc analizados, al menos cuatro transcritos diferentes mostraron que estaban presentes

debido al ajuste alternativo o a la poliadenilación alternativa. El análisis de secuencias de los clones genómicos en comparación con los clones de ADNc reveló la estructura genómica del gen PCA3. Están presentes tres intrones y 4 exones. El primer intrón es aproximadamente de 20 kb de longitud.

5 La primera especie de ADNc se encuentra aproximadamente en un 5% de los clones de ADNc y contiene los exones 1, 2, 3, 4a y 4b (la poliadenilación después de 4b está precedida por una señal de adición de poli-A de consenso real) (figura 1).

10 La segunda especie de ADN, encontrada aproximadamente en un 15% del ADNc clonado, contiene los exones 1, 3, 4a, 4b y 4c, surge por ajuste alternativo del segundo exón (no presente en este ADNc) y termina en una señal de adición poli-a diferente (consenso real) (figura 1).

15 La tercera especie de ADNc contiene los exones 1, 3, 4a y 4b y es la más común encontrada (aproximadamente un 65% de los 80 clones) (figura 1). Este ADNc es muy probablemente responsable de la transcripción más prominente observada por análisis de curvas Northern (2 kb).

20 La cuarta especie de ADNc detectada contiene los exones 1, 3 y 4a que representan aproximadamente un 15% de clones, y termina después de 4a, que es el sitio de detención del clon de DD3 original (figura 1). La señal de adición poli-A presente en el presente documento está próxima a la secuencia consenso.

25 El PCA3 es un gen en el que se produce un ajuste alternativo significativo (así como poliadenilación alternativa), como se evidencia por los transcritos dimensionados observados en las curvas Northern y los diferentes tipos de clones identificados. Como se menciona previamente, se pueden identificar otras variantes de ajuste, ya que virtualmente cualquier combinación de exones es posible. Por ejemplo, se ha identificado recientemente un clon de ADNc que tiene los 2, 3, 4a, 4b y 4c. De hecho, parece que se han aislado clones que representan virtualmente todas las combinaciones de exones posibles.

30 Se identificó una variante de ajuste de este tipo secuenciando un clon denominado λDD3.6. El λDD3.6 es un clon de λgt11 identificado y aislado después de cribar una colección de ADNc realizada a partir de ARN de próstata de un macho de 25 años (obtenido de Clontech) con una sonda de PCA3. El λDD3.6 contiene los exones 3, 4a, 4b, 4c y 4d. Sin embargo, este clon de ADNc también contiene secuencias de intrón (parte del intrón 2, así como el intrón 3).

Una comparación de los dos clones depositados PMB9 y λDD3.6 se muestra en la figura 3.

35 Se examinaron diferentes combinaciones de exones por análisis por ordenador para identificar marcos de lectura abiertos (ORF) y para predecir la región codificante de la proteína. El ORF más largo también fue la región codificante de la proteína más altamente probable. El ORF más largo de 153 nucleótidos codifica un péptido pequeño de 51 aminoácidos, PCA3. El PCA3 está codificado por parte del exón 3 y 4a. El tamaño pequeño de la proteína sugiere que la proteína más probablemente funciona como una molécula mensajera y tiene el potencial para segregarse de las células. La secuencia de nucleótidos de los exones 1-4a-d y la secuencia de aminoácidos de PCA3 se muestran en la fig. 2 fig. 5 (SEQ ID NO:1 y 6 y 2 y 7, respectivamente).

45 Se reconocerá por el experto en la técnica que un clon de ADNc que comprende la secuencia de ácido nucleico presentada en SEQ ID NO:6 y mostrada en la fig. 5 se puede obtener como se describe previamente aislando y caracterizando clones de ADNc de PCA3. Por ejemplo, y como es conocido comúnmente en la técnica, las sondas que son específicas para al menos uno de los extremos 5', exones 1, 2, 3, 4a, 4b, 4c y 4d se puede usar adicionalmente para incrementar la probabilidad de tener un clon de ADNc de PCA3 de longitud completa. Se pueden usar placas de 96 pocillos, por ejemplo, para cribar un gran número de clones de ADNc positivos de PCA3, usando las sondas mencionadas anteriormente. Por supuesto, también se pueden secuenciar clones positivos de PCA3, como es conocido comúnmente y como se describe en el presente documento, hasta que se obtienen un clon de ADN deseado.

55 Además, también es posible obtener un clon de ADNc que comprende la secuencia mostrada en SEQ ID NO:6 y mostrada en la fig. 5, usando cebadores específicos de PCA3 y un método de amplificación tal como PCR. Por ejemplo, se pudo usar tecnología de PCR con cebadores específicos para el extremo 5' y 3' final del ADNc de PCA3, para amplificar un producto deseado (de casi 4 kb) a partir de ARN aislado, por ejemplo, a partir de tumores prostáticos, y clonar los productos de PCR. Sin embargo, puesto que la amplificación por PCR puede introducir errores, se podría requerir más probablemente una secuenciación del ADNc completo.

60 Como es bien conocido por el experto en la técnica, también se puede construir un clon de ADN que comprende la secuencia mostrada en SEQ ID NO:6 y mostrada en la fig. 5 usando los clones descritos en el presente documento (o los aislados nuevamente) y métodos de ingeniería genética convencionales.

65 Por ejemplo, se puede construir un clon de ADNc de longitud completa usando los clones depositados pMB9 y λDD3.6. Un ejemplo no limitante de una estrategia para construir un clon de ADNc de este tipo que comprende la

secuencia de ácido nucleico de SEQ ID NO:6 y fig. 5, se describe a continuación.

Se digirió por completo ADN de fago de λ DD3.6 con NdeI, y se aisló el fragmento NdeI de aproximadamente 2 kb de un gel de agarosa. Este fragmento contiene parte del exón 4b de PCA3, los exones 4c + 4d y aproximadamente 50 nucleótidos de ADN de fago. Después, se rellenan los extremos de este fragmento de 2 kb con ADN polimerasa con fragmento Klenow y dNTP, a continuación se liga el fragmento con extremo romo en los sitios *HincII*/*SmaI* de Bluescript SK. La pérdida del sitio *HindIII* de Bluescript por la digestión de *HincII* y *SmaI* es esencial para las etapas de clonación adicionales en esta estrategia particular (véase a continuación). Cabe señalar que también están presentes sitios NdeI en el fago λ gt11, lo que da lugar a varios fragmentos adicionales, de los que algunos están próximos a 2 kb (es decir, un fragmento de 1,8 kb y uno de 2,5 kb). No obstante, es sencillo separar estas diferentes bandas en un gel de agarosa. La orientación correcta de inserción del fragmento NdeI de 2kb con extremo romo de λ DD3.6 en Bluescript (denominado construcción PCA3-X) se puede verificar por una única digestión de *SacI*, proporcionaría un fragmento de ~0,45 y uno de ~4,5 kb por tinción con bromuro de etidio de geles de agarosa. Se puede realizar el análisis de secuencias para confirmar la identidad del inserto de PCA3.

Después, se digiere por completo la construcción de PCA3-X con *HindIII* y *BamHI* y se aísla un inserto de vector de 4,8 kb a partir de un gel de agarosa. Esto da como resultado la retirada de ~0,2 kb de ADN del inserto. Se puede digerir por completo simultáneamente pMB9 con *BamHI* y *HindIII* y se puede aislar el inserto de 1,9 kb (que contiene los exones de PCA3 1, 2, 3, 4a y la mayoría del exón 4b) a partir de un gel de agarosa. El inserto derivado de pMB9 se liga en el sitio *BamHI*/*HindIII* de la construcción de PCA3-X. La construcción resultante, PCA3-Y contiene el ADNc completo de PCA3, excepto por los 22 primeros nucleótidos del exón 1 (véase a continuación y en la fig. 4). Se pueden añadir estos 22 nucleótidos al ADNc de PCA3 cortando la construcción PCA3-Y y el oligó-(74)-mero (SEQ ID NO:8) por completo con *BamHI* y *PstI* y ligando el oligómero en la construcción PCA3-Y, lo que da como resultado la construcción PCA3-Z. Se pueden realizar análisis de secuencias de nucleótidos para verificar que el oligo se ligó apropiadamente (es decir, para confirmar que sólo se ligó un oligo y no todo el conjunto de oligos). Por supuesto, se puede realizar una secuenciación del ADNc resultante en PCA3-Z, para verificar la integridad de la secuencia de ácido nucleico.

El cribado de un panel híbrido de células somáticas reveló que el gen que codifica PCA3 está situado en el cromosoma 9 humano. Usando una mezcla de cuatro clones genómicos relacionados con PCA3 como sonda para hibridar a cromosomas de la metafase de linfocitos humanos, se mapeó el PCA3 para 9q21-22 (véase también la figura 1).

Se estudió la conservación del gen PCA3 durante la evolución por análisis de bandas Southern y reveló que un homólogo de este gen está presente en monos, vacas, caballos, ovejas, cabras y cerdos. El gen también está presente en perros y gatos. Por comparación, el gen que codifica PSA sólo se encuentra en primates.

Ejemplo 2

Expresión específica de próstata de PCA3

Después del desarrollo de cebadores específicos de PCA3, se realizó un análisis de RT-PCR usando ARN de varios tejidos humanos normales. A 40 ciclos de PCR, se amplificaron productos relacionados con PCA3 en tejidos de BPH y próstata normales. La expresión de PCA3 es muy específica de próstata puesto que no se pudo amplificar ningún producto de PCA3 bajo estas condiciones en los siguientes tejidos humanos normales: arteria, cerebro, mama, vejiga, colon, duodeno, corazón, hígado, pulmón, ovario, páncreas, placenta, vesículas seminales, músculo esquelético, piel, espina dorsal, bazo y testículos. Además, en las líneas celulares de cáncer de próstata ALVA-31, DU145, JCA-1, PPC-1, PC3, y TSU-Pr1 no se pudo detectar producto de PCR relacionado con PCA3. En la línea celular LNCaP se puede obtener un producto después de 40 ciclos de PCR (mientras que bajo las mismas condiciones se puede obtener un producto en tumores prostáticos en 20 ciclos). La tecnología usada para evaluar la expresión específica prostática de PCA3 se puede adaptar en una prueba de diagnóstico para el cáncer de próstata. Además, se puede adaptar para la identificación del origen prostático de una metástasis.

Además, se realizó un análisis de RT-PCR semicuantitativa para comparar la expresión de PCA3 con la de PSA (antígeno específico de próstata) y PSM (antígeno de membrana específico de próstata) y para establecer si se puede usar un análisis de PCA3-RT-PCR para distinguir especímenes prostáticos malignos de benignos. Después de la cuantificación de la reacción RT, se usaron 10 ng de ADNc para la reacción de PCR y también se examinó como control, beta-2 microglobulina. Los productos de PCA3 encontrados permitieron una clara distinción entre especímenes benignos y and malignos en 23 de 25 casos estudiados mientras que PSA y PSM no pudieron hacer esta distinción: se encontraron cantidades aproximadamente iguales de producto en muestras normal y de tumor. También se comparó la expresión de PSA y PCA3 por análisis de bandas de Northern, lo que mostró claramente la mayor especificidad de tumor de PCA3. Se observó al menos una sobreexpresión de 20 veces de PCA3 en carcinomas prostáticos en comparación con tejidos normal o BPH. Esto es distintivamente diferente de la expresión de PSM y PSA, ambos disminuyen en tejidos malignos frente a benignos. Por tanto, parece que PCA3 es un buen marcador para el diagnóstico de cáncer de próstata.

Un marcador de tumor ideal para el cáncer de próstata no debe sólo poder distinguir positivamente entre tejidos benignos y malignos sino que también deber poder predecir la evolución clínica (cura o progresión) de pacientes afligidos con esta enfermedad. Los datos han mostrado que, de hecho, el nivel de expresión de PCA3 tiene a correlacionarse positivamente con el grado de tumor.

Se usa RISH (y finalmente inmunohistoquímica) para establecer si existe o no correlación entre la sobreexpresión de PCA3, grado de tumor, fase y evolución clínica. Para especímenes tanto incrustados en parafina como congelados, está disponible un seguimiento clínico a largo plazo. Usando análisis por imagen asistida por ordenador, se realiza la cuantificación de los niveles de expresión de PCA3 detectados por RISH y esto se normaliza a una referencia externa (Tamimi et al., Cancer Res. 53: 5512-16 (1993); Tamimi et al., B.J. Cancer (1996)). Se usará un análisis de regresión multivariable que incluye grado de Gleason, fase de tumor patológico, fase de tumor clínico, niveles de PSA y expresión de PCA3, para establecer si PCA3 es un factor predictivo exacto de la progresión y tiene valor de pronóstico (adicional).

Se han desarrollado ensayos de reacción en cadena de la polimerasa y transcriptasa inversa para detectar células micrometastásicas hematógenas ocultas que de otro modo podrían haber pasado sin detectarse por modalidades de estadificación actualmente disponibles. Dichos ensayos de RT-PCR ya se han realizado en pacientes con cáncer de próstata y otras neoplasias malignas. Se puede usar un ensayo de RT-PCR altamente sensible (con cebadores internos) (u otros tipos de ensayos de amplificación incluidos, pero sin limitación, NASBA, PCR, QB rep., SOA, TMA y LCR (Winn-Deen, J. Clin., Liquid Assay 19: 21-26 (1996)) para detectar células de cáncer de próstata en la sangre en circulación de pacientes con cáncer de próstata para identificar pacientes con riesgo de tener o desarrollar metástasis. Los experimentos incluirán controles apropiados (por ejemplo, β -2-microglobulina) y se realizarán de forma semicuantitativa (es decir, cuantificar la síntesis de ADNc y usar cantidades iguales de entrada para el análisis de PCR).

Los estudios de estadificación molecular se realizarán en el contexto más amplio del programa BIOMED II (*Markers for Prostate Cancer*). En este estudio de colaboración extenso, se estudiarán PSA y PSM así como otros marcadores potencialmente interesantes para el cáncer de próstata. Las muestras de sangre ya se han recogido de los pacientes que se les ha diagnosticado con enfermedad prostática en las instituciones participantes. Se ha iniciado una optimización de la recogida y el manejo de muestras de sangre de pacientes para la detección de células tumorales en circulación. El uso, por ejemplo, de tubos CPT de vacutainer™ (Beckton-Dickinson) para la recogida de sangre y la purificación de leucocitos de sangre periférica en combinación con un procedimiento de extracción de ARN Trizol™ (a base de tiocianato de guanidinio) dio como resultado la preparación de ARN cualitativa y cuantitativamente adecuado para el análisis de PCR. El uso de cebadores específicos de PCA3 para amplificar transcritos de PCA3 en ARN extraído de sangre de pacientes con cáncer de próstata reveló que se podía detectar la presencia de células de cáncer de próstata en la circulación sanguínea, no sólo de pacientes con metástasis probada, sino también de pacientes con enfermedad localizada supuesta. Se llevarán a cabo estudios más extensos sobre una población de pacientes más grande y una correlación con datos clínicos y seguimiento para determinar el valor de pronóstico de PCA3 para el paciente con cáncer de próstata individual.

Los análisis de RT-PCR con cebadores internos (o métodos de amplificación similares) deben resultar fundamentales en la determinación de si hay algún órgano (aún no sometido a prueba) que exprese PCA3. Por ejemplo, la glándula de Cowper (mismo origen embrionario que la próstata) y también la glándula de Skene ("homólogo" femenino a la próstata) se someterán a prueba para PCA3.

En un espécimen de tejido de próstata "normal" que contiene un 10% de células tumorales, se detectó expresión de PCA3 lo que indica la alta sensibilidad de PCA3 como marcador tumoral. De esta manera, también se detectó la expresión de PCA3 en unas pocas muestras de BPH que se encontró posteriormente que contenían pequeñas áreas de células tumorales. El nivel de expresión de PCA3 en cánceres prostáticos muestra una tendencia hacia una correlación positiva con grado tumoral. Estos datos se basan en el análisis de autorradiografías resultantes de la hibridación de curvas Northern.

La observación de que la expresión de PCA3 parece que se incrementa con la pérdida de la diferenciación es diferente de los informado para PSA, puesto que los niveles de expresión de PSA disminuyen con la pérdida de la diferenciación (Hakalahti et al., Int. J. Cancer 55:590-597 (1993)). Existe al menos una sobreexpresión de 20 veces de PCA3 en carcinomas prostáticos en comparación con tejidos normal o BPH. Esto es distintivamente diferente de la expresión de PSA que informa de disminuir en tejidos malignos frente a benignos. Se detectó expresión de PCA3 en 4 de 4 metástasis estudiadas.

Ejemplo 3

Identificación de un sitio de inicio de la transcripción de PCA3

Para determinar el sitio de inicio de la transcripción de PCA3 se realizaron análisis de extensión de cebador, mapeo de nucleasa S1 y ensayos de RACE 5' (amplificación rápida de extremos de ADNc). Se encontró que el sitio de inicio de la transcripción principal estaba situado dentro de un intervalo de de 4 nucleótidos (figura 4).

Los resultados de estos experimentos alargan adicionalmente el tamaño del ADNc en sentido 5' por 22 nt adicionales con respecto a la secuencia de ADNc de pMB9 (SEQ ID NO:1 y figura 2). Esta secuencia de polinucleótidos 5' adicional también se muestra en SEQ ID NO:6 y la figura 5).

5

Listado de secuencias

<110> Bussemakers, Marion J. G.
Isaacs, William B.

10 <120> PCA3, genes PCA3 y métodos de uso

<130> K3129 EP/1 S3

<150> documento EP 9896696.2

<151> 1998-04-09

<150> documento PCT/CA98/00346

15 <151> 1998-04-09

<150> documento US 60/041.836

<151> 10-04-1997

<160> 9

<170> PatentIn versión 3.5

20

<210> 1

<211> 2037

<212> ADN

<213> Homo sapiens

25

<220>

<221> CDS

<222> (379)...(531)

30

<220>

<221> característica_misc.

<222> (1472)...(1472)

<223> n = a, t, c o g

35

<220>

<221> característica_misc.

<222> (1517)...(1517)

<223> n = a, t, c o g

40

<220>

<221> característica_misc.

<222> (1563)...(1563)

<223> n = a, t, c o g

45

<400> 1

agaagctggc atcagaaaaa cagaggggag atttgtgtgg ctgcagccga gggagaccag 60

gaagatctgc atggtgggaa ggacctgatg atacagagga attacaacac atatacttag 120

tgtttcaatg aacaccaaga taaataagtg aagagctagt ccgctgtgag tctcctcagt 180

gacacagggc tggatcacca tcgacggcac tttctgagta ctgagtgcag caaagaaaga 240

ctacagacat ctcaatggca ggggtgagaa ataagaaagg ctgctgactt taccatctga 300

ggccacacat ctgctgaaat ggagataatt aacatcacta gaaacagcaa gatgacaata 360

taatgtctaa gtagtgac atg ttt ttg cac att tcc agc ccc ttt aaa tat	411
Met Phe Leu His Ile Ser Ser Pro Phe Lys Tyr	
1 5 10	
cca cac aca cag gaa gca caa aag gaa gca cag aga tcc ctg gga gaa	459
Pro His Thr Gln Glu Ala Gln Lys Glu Ala Gln Arg Ser Leu Gly Glu	
15 20 25	
atg ccc ggc cgc cat ctt ggg tca tgc atg agc ctc gcc ctg tgc ctg	507
Met Pro Gly Arg His Leu Gly Ser Ser Met Ser Leu Ala Leu Cys Leu	
30 35 40	
gtc ccg ctt gtg agg gaa gga cat tagaaaaatga attgatgtgt tccttaaagg	561
Val Pro Leu Val Arg Glu Gly His	
45 50	
atgggcagga aaacagatcc tgttgtggat atttatttga acgggattac agatttgaaa	621
tgaagtcaca aagtgagcat taccaatgag agggaaaacag acgagaaaaat cttgatggct	681
tcacaagaca tgcaacaaac aaaatggaat actgtgatga catgaggcag ccaagctggg	741
gaggagataa ccacggggca gagggtcagg attctggccc tgcctgcctaa actgtgcgtt	801
cataacaaaa tcatttcata tttctaacc tcaaaacaaa gctgttgtaa tatctgatct	861
ctacggttcc ttctgggccc aacattctcc atatatccag ccacactcat ttttaatat	921
tagttcccag atctgtactg tgacctttct acactgtaga ataacattac tcattttgtt	981
caaagaccct tcgtgttgct gcctaatatg tagctgactg tttttcctaa ggagtgttct	1041
ggcccagggg atctgtgaac aggctgggaa gcatctcaag atctttccag gggtatactt	1101
actagcacac agcatgatca ttacggagtg aattatctaa tcaacatcat cctcagtgtc	1161
tttgcccata ctgaaattca tttcccactt ttgtgcccat tctcaagacc tcaaaatgtc	1221
attccattaa tatcacagga ttaacttttt tttttaacct ggaagaattc aatgttacat	1281
gcagctatgg gaatttaatt acatattttg ttttccagtg caaagatgac taagtccttt	1341
atccctcccc ttgttttgat tttttttcca gtataaagtt aaaatgctta gccttgact	1401
gaggctgtat acagcacagc ctctcccat cctccagcc ttatctgtca tcaccatcaa	1461
cccctcccat nysacctaaa caaatctaa cttgtaattc cttgaacatg tcaggncata	1521
catrttctct tctgcctgag aagetcttcc ttgtctctta antctagaat gatgtaaagt	1581
tttgaataag ttgactatct tacttcatgc aaagaaggga cacatatgag attcatcatc	1641
acatgagaca gcaaatacta aaagtgtaat ttgattataa gagtttagat aaatatatga	1701
aatgcaagak ccacagaggg aatgtttatg gggcacgttt gtaagcctgg gatgtgaagm	1761
aaaggcaggg aacctcatag tatcttatat aatatacttc atttctctat ctctatcaca	1821
atattcaaca agctttttcac agaattcatg cagtgc aaat ccccaaagggt aacctttatc	1881
catttcatgg tgagtgcgct ttagaatttt ggcaaatcat actggctact tatctcaact	1941
ttgagatgtg tttgtccttg tagttaattg aaagaaatag ggcactcttg tgagccactt	2001
tagggttcac tcttggaat aaagaattta caaaga	2037

ES 2 402 947 T3

<211> 51
 <212> PRT
 <213> Homo sapiens

5 <400> 2

Met Phe Leu His Ile Ser Ser Pro Phe Lys Tyr Pro His Thr Gln Glu
 1 5 10 15

Ala Gln Lys Glu Ala Gln Arg Ser Leu Gly Glu Met Pro Gly Arg His
 20 25 30

Leu Gly Ser Ser Met Ser Leu Ala Leu Cys Leu Val Pro Leu Val Arg
 35 40 45

Glu Gly His
 50

10 <210> 3
 <211> 1872
 <212> ADN
 <213> Homo sapiens

15 <220>
 <221> característica_misc.
 <222> (1307)...(1307)
 <223> n = a, t, c o g

20 <220>
 <221> característica_misc.
 <222> (1352)...(1352)
 <223> n = a, t, c o g

25 <220>
 <221> característica_misc.
 <222> (1398)...(1398)
 <223> n = a, t, c o g

<400> 3

agaagctggc atcagaaaaa cagaggggag atttgtgtgg ctgcagccga gggagaccag 60
 gaagatctgc atggtgggaa ggacctgatg atacagaggt gagaaataag aaaggctgct 120
 gactttacca tctgaggcca cacatctgct gaaatggaga taattaacat cactagaaac 180

ES 2 402 947 T3

agcaagatga caatataatg tctaagtagt gacatgtttt tgcacatttc cagccccctt	240
aaatatccac acacacagga agcacaaaag gaagcacaga gatccctggg agaaatgccc	300
ggccgccatc ttgggtcadc gatgagcctc gccctgtgcc tggccccgct tgtgaggga	360
ggacattaga aaatgaattg atgtgttctt taaaggatgg gcaggaaaac agatcctgtt	420
gtggatattt atttgaacgg gattacagat ttgaaatgaa gtcacaaagt gagcattacc	480
aatgagagga aaacagacga gaaaatcttg atggcttcac aagacatgca acaaacaaaa	540
tggaatactg tgatgacatg aggcagccaa gctggggagg agataaccac ggggcagagg	600
gtcaggattc tggccctgct gcctaaactg tgcgttcata accaaatcat ttcataattc	660
taaccctcaa aacaaagctg ttgtaatatc tgatctctac ggttccttct gggcccaaca	720
ttctccatat atccagccac actcattttt aatatttagt tcccagatct gtactgtgac	780
cttctacac tgtagaataa cattactcat tttgttcaaa gacccttcgt gttgctgcct	840
aatatgtagc tgactgtttt tcctaaggag tgttctggcc caggggatct gtgaacaggc	900
tgggaagcat ctcaagatct ttccagggtt atacttacta gcacacagca tgatcattac	960
ggagtgaatt atctaataa catcatcctc agtgtctttg ccataactga aattcatttc	1020
ccacttttgt gccattctc aagacctcaa aatgtcattc cattaatatc acaggattaa	1080
cttttttttt taacctggaa gaattcaatg ttacatgcag ctatgggaat ttaattacat	1140
attttgtttt ccagtgcata gatgactaag tcctttatcc ccccccttg tttgattttt	1200
tttccagtat aaagttaaaa tgcttagcct tgtactgagg ctgtatacag cacagcctct	1260
ccccatccct ccagccttat ctgtcatcac catcaacccc tcccatnysa cctaaacaaa	1320
atctaacttg taattccttg aacatgtcag gncatacatt rtcccttctg cctgagaagc	1380
tcttccctgt ctcttaantc tagaatgatg taaagttttg aataagttga ctatcttact	1440
tcagtcaaag aaggagacac tatgagattc atcatcacat gagacagcaa atactaaaag	1500
tgaatttga ttataagagt ttagataaat atatgaaatg caagakccac agagggaatg	1560
tttatggggc acgtttgtaa gcctgggatg tgaagmaaag gcagggaacc tcatagtatc	1620
ttatataata tacttcattt ctctatctct atcacaatat ccaacaagct ttacacagaa	1680
ttcatgcagt gcaaatcccc aaaggttaacc tttatccatt tcatgggtgag tgcgctttag	1740
aattttggca aatcatactg gtcacttata tcaactttga gatgtgtttg tccttgtagt	1800
taattgaaag aaatagggca ctcttgtagg ccactttagg gttcactcct ggcaataaag	1860
aatttacaaa ga	1872

<210> 4

<211> 820

5 <212> ADN

<213> Homo sapiens

<400> 4

ES 2 402 947 T3

```

agaagctggc atcagaaaaa cagaggggag atttgtgtgg ctgcagccga gggagaccag      60
gaagatctgc atggtgggaa ggacctgatg atacagaggt gagaaataag aaaggctgct      120
gactttacca tctgaggcca cacatctgct gaaatggaga taattaacat cactagaaac      180
agcaagatga caatataatg tctaagtagt gacatgtttt tgcacatttc cagccccctt      240
aaatatccac acacacagga agcacaaaag gaagcacaga gatccctggg agaaatgccc      300
ggccgccatc ttgggtcatc gatgagcctc gccctgtgcc tggteccgct tgtgagggaa      360
ggacattaga aaatgaattg atgtgttctt taaaggatgg gcaggaaaac agatcctggt      420
gtggatattt atttgaacgg gattacagat ttgaaatgaa gtcacaaagt gagcattacc      480
aatgagagga aaacagacga gaaaatcttg atggcttcac aagacatgca acaacaaaaa      540
tggaatactg tgatgacatg aggcagccaa gctggggagg agataaccac ggggcagagg      600
gtcaggattc tggccctgct gcctaaactg tgcgttcata accaaatcat ttcataattc      660
taacctcaa aacaaagctg ttgtaatatc tgatctctac gggtccttct gggcccaaca      720
ttctccatat atccagccac actcattttt aatatttagt tcccagatct gtactgtgac      780
ctttctacac tgtagaataa cattactcat tttgttcaaa      820

```

<210> 5
 <211> 11
 5 <212> PRT
 <213> Homo sapiens

<400> 5
 His Thr Gln Glu Ala Gln Lys Glu Ala Gln Arg
 1 5 10

10 <210> 6
 <211> 3582
 <212> ADN
 <213> Homo sapiens

15 <220>
 <221> CDS
 <222> (401)...(553)

20 <400> 6
 acagaagaaa tagcaagtgc cgagaagctg gcatcagaaa aacagagggg agatttgtgt 60

ES 2 402 947 T3

ggctgcagcc gagggagacc aggaagatct gcatggtggg aaggacctga tgatacagag	120
gaattacaac acatatactt agtgtttcaa tgaacaccaa gataaataag tgaagagcta	180
gtccgctgtg agtctcctca gtgacacagg gctggatcac catcgacggc actttctgag	240
tactcagtgc agcaaagaaa gactacagac atctcaatgg caggggtgag aaataagaaa	300
ggctgctgac ttaccatct gaggccacac atctgctgaa atggagataa ttaacatcac	360
tagaaacagc aagatgacaa tataatgtct aagtagtgac atg ttt ttg cac att	415
Met Phe Leu His Ile	
1 5	
tcc agc ccc ttt aaa tat cca cac aca cag gaa gca caa aag gaa gca	463
Ser Ser Pro Phe Lys Tyr Pro His Thr Gln Glu Ala Gln Lys Glu Ala	
10 15 20	
cag aga tcc ctg gga gaa atg ccc ggc cgc cat ctt ggg tca tgc atg	511
Gln Arg Ser Leu Gly Glu Met Pro Gly Arg His Leu Gly Ser Ser Met	
25 30 35	
agc ctc gcc ctg tgc ctg gtc ccg ctt gtg agg gaa gga cat	553
Ser Leu Ala Leu Cys Leu Val Pro Leu Val Arg Glu Gly His	
40 45 50	
tagaaaatga attgatgtgt tccttaaagg atgggcagga aaacagatcc tgttgtggat	613
atattattga acgggattac agatttgaaa tgaagtcaca aagtgagcat taccaatgag	673
aggaaaacag acgagaaaat cttgatggct tcacaagaca tgcaacaaac aaaatggaat	733
actgtgatga catgaggcag ccaagctggg gaggagataa ccacggggca gagggtcagg	793
attctggccc tgcctcctaa actgtgcgtt cataaccaa tcatttcata tttctaacc	853
tcaaaacaaa gctgtttaa tatctgatct ctacggttcc ttctgggccc aacattctcc	913
atatatccag ccacactcat ttttaatat tagttcccag atctgtactg tgacctttct	973
acactgtaga ataacattac tcattttgtt caaagacctc tcgtgttgc gctaatatg	1033
tagctgactg tttttcctaa ggagtgttct ggcccagggg atctgtgaac aggtctggaa	1093
gcatctcaag atctttccag ggttatactt actagcacac agcatgatca ttacggagtg	1153
aattatctaa tcaacatcat cctcagtgtc ttgcccata ctgaaattca tttcccactt	1213
ttgtgcccac tctcaagacc tcaaaatgtc attccattaa taccacagga ttaacttttt	1273
tttttaacct ggaagaattc aatgttacat gcagctatgg gaatttaatt acatattttg	1333
ttttccagtg caaagatgac taagtccttt atccctcccc ttgtttgat tttttttcca	1393
gtataaagtt aaaatgctta gccttgact gaggtgtat acagcacagc ctctcccat	1453
ccctccagcc ttatctgtca tcaccatcaa cccctcccat accacctaaa caaaatctaa	1513

cttgtaattc cttgaacatg tcaggacata cattattcct tctgcctgag aagctcttcc	1573
ttgtctctta aatctagaat gatgtaaagt tttgaataag ttgactatct tacttcatgc	1633
aaagaaggga cacatatgag attcatcatc acatgagaca gcaaatacta aaagtgtaat	1693
ttgattataa gagtttagat aaatatatga aatgcaagag ccacagaggg aatgtttatg	1753
gggcacgttt gtaagcctgg gatgtgaagc aaaggcaggg aacctcatag tatcttatat	1813
aatatacttc atttctctat ctctatcaca atateccaaca agcttttcac agaattcatg	1873
cagtgc aaat ccccaaaggt aacctttatc catttcatgg tgagtgcgct ttagaatttt	1933
ggcaaatcat actggctact tatctcaact ttgagatgtg tttgtccttg tagttaattg	1993
aaagaaatag ggcactcttg tgagccactt tagggttcac tcctggcaat aaagaattta	2053
caaagagcta ctcaggacca gttgttaaga gctctgtgtg tgtgtgtgtg tgtgtgtgag	2113
tgtacatgcc aaagtgtgcc tctctctctt gaccattat ttcagactta aaacaagcat	2173
gttttcaaat ggcactatga gctgccaatg atgtatcacc accatatctc attattctcc	2233
agtaaagtgt ataataatgt catctgttaa cataaaaaaa gtttgacttc acaaaagcag	2293
ctggaaatgg acaaccacaa tatgcataaa tctaactcct accatcagct acacactgct	2353
tgacatatat tgtagaagc acctcgcat tgtgggttct cttaagcaaa atacttgcac	2413
taggtctcag ctggggctgt gcatcaggcg gtttgagaaa tattcaattc tcagcagaag	2473
ccagaatttg aattccctca tcttttagga atcatttacc aggtttggag aggattcaga	2533
cagctcaggt gctttcacta atgtctctga acttctgtcc ctctttgtgt tcatggatag	2593
tccaataaat aatgttatct ttgaactgat gctcatagga gagaatataa gaactctgag	2653
tgatatcaac attagggatt caaagaaata ttagatttaa gctcacactg gtcaaaagga	2713
accaagatac aaagaactct gagctgtcat cgtccccatc tctgtgagcc acaaccaaca	2773
gcaggacca acgcatgtct gagatcctta aatcaaggaa accagtgtca tgagttgaat	2833
tctctatta tggatgctag cttctggcca tctctggctc tcctcttgac acatattagc	2893
ttctagcctt tgcctccacg acttttatct tttctccaac acatcgctta ccaatcctct	2953
ctctgctctg ttgctttgga cttccccaca agaatttcaa cgactctcaa gtcttttctt	3013
ccatccccac cactaacctg aattgcctag acccttattt ttatttaattt ccaatagatg	3073
ctgcctatgg gctaataattg ctttagatga acattagata tttaaagtct aagaggttca	3133
aaatccaact cattatcttc tctttcttcc acctccctg ctctctctcc tatattactg	3193
attgactgaa caggatggct cccaagatgc cagtcaaag agaaaccag tggctccttg	3253
tggatcatgc atgcaagact gctgaagcca gaggatgact gattacgcct catgggtgga	3313

ES 2 402 947 T3

```

ggggaccact cctgggcctt cgtgattgtc aggagcaaga cctgagatgc tccctgcctt 3373
cagtgtcctc tgcattctcc ctttctaattg aagatccata gaatttgcta catttgagaa 3433
ttccaattag gaactcacat gttttatctg cccatcaat tttttaaaact tgctgaaaat 3493
taagtttttt caaaatctgt ccttgtaaatt tactttttct tacagtgtct tggcatacta 3553
tatcaacttt gattctttgt tacaacttt 3582

```

<210> 7
 <211> 51
 5 <212> PRT
 <213> Homo sapiens

<400> 7
 Met Phe Leu His Ile Ser Ser Pro Phe Lys Tyr Pro His Thr Gln Glu
 1 5 10 15
 Ala Gln Lys Glu Ala Gln Arg Ser Leu Gly Glu Met Pro Gly Arg His
 20 25 30
 Leu Gly Ser Ser Met Ser Leu Ala Leu Cys Leu Val Pro Leu Val Arg
 35 40 45
 Glu Gly His
 50

10 <210> 8
 <211> 74
 <212> ADN
 <213> Homo sapiens

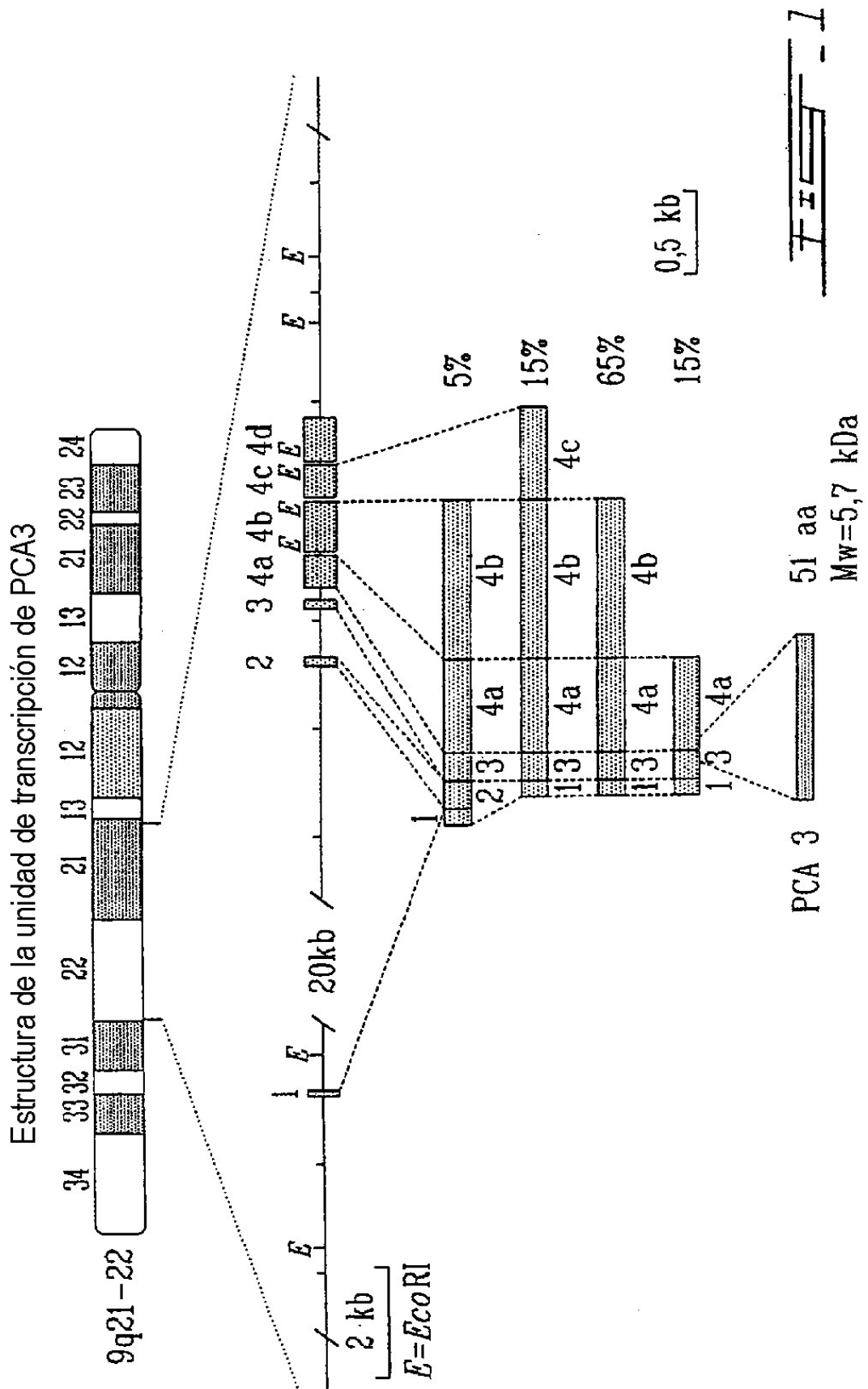
15 <400> 8
 ggatccacag aagaaatagc aagtgccgag aagctggcat cagaaaaaca gaggggagat 60
 ttgtgtggct gcag 74

20 <210> 9
 <211> 123
 <212> ADN
 <213> Homo sapiens

<400> 9
 gacagaagaa atagcaagtg ccgagaagct ggcattcagaa aaacagaggg gagatttgtg 60
 tggctgcagc cgagggagac caggaagatc tgcattggtg gaaggacctg atgatacaga 120
 25 ggt 123

REIVINDICACIONES

1. Un método para determinar si un ácido nucleico de PCA3 está presente en una muestra que comprende ARN, mostrando dicho ácido nucleico de PCA3 sobreexpresión en células de cáncer de próstata en relación con células de próstata normales o tejidos de próstata benignos, comprendiendo dicho método las etapas de:
 - (a) sintetizar ADNc usando ARN de la muestra de prueba como molde; y
 - (b) detectar entre el ADNc sintetizado un ácido nucleico que comprende un marco de lectura abierto (ORF) que codifica SEQ ID NO: 7, en el que la Arg23 del ORF está codificada por el codón AGA que resulta del ajuste del ARN del exón 3 del PCA3 al exón 4a del PCA3, determinando por ello que el ácido nucleico de PCA3 está presente en la muestra de prueba.
2. El método de la reivindicación 1, en el que la etapa (a) comprende alinear un cebador sintético adyacente a una secuencia diana contenida dentro del ARN de la muestra de prueba, en el que dicho cebador sintético es complementario a una secuencia contenida dentro de un ácido nucleico que comprende el exón 3 del PCA3 y el exón 4a del PCA3.
3. El método de la reivindicación 1 o 2, en el que el ácido nucleico de PCA3 comprende la secuencia del ORF que se extiende desde la posición de base 401 a la 553 de SEQ ID NO: 6.
4. El método de una cualquiera de las reivindicaciones 1 a 3, en el que la muestra de prueba es un extracto de ácido nucleico de células o fluidos biológicos.
5. El método de una cualquiera de las reivindicaciones 1 a 3, en el que la muestra de prueba es una muestra biológica seleccionada del grupo que consiste en un organismo completo, una muestra de órgano, una muestra de tejido, una muestra de sangre y una muestra de células.
6. El método de una cualquiera de las reivindicaciones 2 a 5, en el que el cebador sintético tiene una longitud de:
 - (i) desde 10 hasta 50 nucleótidos;
 - (ii) desde 10 hasta 35 nucleótidos; o
 - (iii) desde 20 hasta 35 nucleótidos.
7. El método de una cualquiera de las reivindicaciones 1 a 6, en el que dicho cebador sintético y ARN de la muestra de prueba pueden formar una estructura bicatenaria en una solución de formamida al 50 %, SSC 5x, solución de Denhardt 5x, SDS al 1 % y 100 µg/ml de ADN de esperma de salmón desnaturalizado a 68 °C.
8. El método de la reivindicación 6, en el que el cebador sintético es complementario a una secuencia contenida dentro del exón 4a.
9. El método de la reivindicación 6, en el que el cebador sintético comprende una secuencia contenida dentro del exón 3.
10. El método de una cualquiera de las reivindicaciones 2 a 9, en el que la etapa (a) de alineación comprende combinar ARN de la muestra de prueba con el cebador sintético en una mezcla de reacción para realizar la RT-PCR.
11. El método de una cualquiera de las reivindicaciones 1 a 10, que comprende además una etapa para determinar si el ácido nucleico de PCA3 que se determinó que estaba presente en la muestra de prueba está presente en una cantidad mayor que la cantidad presente en una muestra que comprende ácidos nucleicos aislados a partir de células de tejido de próstata normal.
12. El método de la reivindicación 11, que comprende además una etapa para detectar un ácido nucleico que codifica el PSA (antígeno específico de próstata) o un ácido nucleico que codifica el PSM (antígeno de membrana específico de próstata).
13. El método de una cualquiera de las reivindicaciones 1 a 12, en el que la etapa (b) de detección comprende detectar la secuencia de ácido nucleico que codifica el ORF de PCA3 completo.



```

1      99      379      531
1      98      264      447      986
          263      446      985      2037
=====//=====//=====
>-----EXON1->-----EXON2->-----EXON3->-----EXON4a--//-----EXON4b--//----->
1 |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
   >-----PCA3----->
                                     FEF-2A

```

1) EXON1, 2) EXON2, 3) EXON3, 4) EXON4A, 5) EXON4b, PCA3 (aminoácidos)
 (las señales de adición de poli-A están subrayadas)
 CON) : Intervalo = 1 a 2037

```

Nombre Base
1 1      agaagctggc atcagaaaaa cagaggggag atttgtgtgg ctgcagccga gggagaccag

CON 1      AGAAGCTGGC ATCAGAAAAA CAGAGGGGAG ATTTGTGTGG CTGCAGCCGA GGGAGACCAG

1 61      gaagatctgc atggtgggaa ggacctgatg atacagag
2 1      ga attacaacac atatacttag
                                     FEF-2B

```

CON 61 GAAGATCTGC ATGGTGGGAA GGACCTGATG ATACAGAGGA ATTACAACAC ATATACTTAG
 2 23 tgtttcaatg aacaccaaga taaataagtg aagagctagt ccgctgtgag tctcctcagt
 CON 121 TGTTC AATG AACACCAAGA TAAATAAGTG AAGAGCTAGT CCGCTGTGAG TCTCCTCAGT
 2 83 gacacagggc tggatcacca tcgacggcac tttctgagta ctcagtgag caaagaaag
 CON 181 GACACAGGGC TGGATCACCA TCGACGGCAC TTTCTGAGTA CTCAGTGCAG CAAAGAAAG
 2 142 actacagaca tctcaatggc aggg
 3 1 gtgaga aataagaaag gctgctgact ttaccatctg
 CON 240 ACTACAGACA TCTCAATGGC AGGGGTGAGA AATAAGAAAG GCTGCTGACT TTACCATCTG
 3 37 aggccacaca tctgctgaaa tggagataat taacatcact agaaacagca agatgacaat
 CON 300 AGGCCACACA TCTGCTGAAA TGGAGATAAT TAACATCACT AGAAACAGCA AGATGACAAT
FILE - 2L

```

97   ataatgtcta agtagtgac atgttttttg cacattttcc agcccccttt aaatatcca cacaca

:ON 360   ATAATGTCTA AGTAGTGAC ATGTTTTTIG CACATTTTCC AGCCCCCTTT AAATATCCA CACACA
:CA3 1     M F L H I S S P F K Y P H T

      158   caggaagca caaaaggaa gcacagag
      :      a tccctggga gaaatgccc ggcggccat cttggg

:ON 421   CAGGAAGCA CAAAAGGAA GCACAGAGA TCCCTGGGA GAAATGCCC GGCCGCCAT CTTGGG
:CA3 15   Q E A Q K E A Q R S L G E M P G R H L G

      35   tcatcgatg agcctcgcc ctgtgcctg gtcccgctt gtgAGGGA GGACATTAG AAA

:ON 481   TCATCGATG AGCCTCGCC CTGTGCCTG GTCCCGCTT GTGAGGGAA GGACATTAG AAA
:CA3 35   S S M S L A L C L V P L V R E G H ***

      93   ATGAATTGAT GTGTTTCCTTA AAGGATGGC AGGAAAACAG ATCCTGTGTGT GGATATTAT

:ON 538   ATGAATTGAT GTGTTTCCTTA AAGGATGGC AGGAAAACAG ATCCTGTGTGT GGATATTAT

```

FIG - 20

4 153 TTGAACGGGA TTACAGATTT GAAATGAAGT CACAAAGTGA GCATTACCAA TGAGAGGAAA
 CON 598 TTGAACGGGA TTACAGATTT GAAATGAAGT CACAAAGTGA GCATTACCAA TGAGAGGAAA
 4 213 ACAGACGAGA AAATCTTGAT GGCTTCACAA GACATGCAAC AAACAAAATG GAATACTGTG
 CON 658 ACAGACGAGA AAATCTTGAT GGCTTCACAA GACATGCAAC AAACAAAATG GAATACTGTG
 4 273 ATGACATGAG GCAGCCAAGC TGGGGaggag ataaccacgg ggcagAGGGT CAGGATTCTG
 CON 718 ATGACATGAG GCAGCCAAGC TGGGGAGGAG ATAACCACGG GGCAGAGGGT CAGGATTCTG
 4 333 GCCCTGCTGC CTAAACTGTG CGTTCATAAC CAAATCATTT CATATTTCTA ACCCTCAAAA
 CON 778 GCCCTGCTGC CTAAACTGTG CGTTCATAAC CAAATCATTT CATATTTCTA ACCCTCAAAA

~~F E S - 2 E~~

4 393 CAAAGCTGTT GTAATATCTG ATCTCTACGG TTCCCTTCTGG GCCCAACATT CTCCATATAT
 CON 838 CAAAGCTGTT GTAATATCTG ATCTCTACGG TTCCCTTCTGG GCCCAACATT CTCCATATAT
 4 453 CCAGCCACAC TCATTTTAA TATTTAGTTC CCAGATCTGT ACTGTGACCT TTCTACACTG
 CON 898 CCAGCCACAC TCATTTTAA TATTTAGTTC CCAGATCTGT ACTGTGACCT TTCTACACTG
 4 513 TAGAATAACA TTACTCATTT TGTTCAAA
 5 1 GA CCTTTCGTGT TGCTGCCCTAA TATGTAGCTG
 CON 958 TAGAATAACA TTACTCATTT TGTTCAAA GA CCTTTCGTGT TGCTGCCCTAA TATGTAGCTG
 5 33 ACTGTTTTTC CTAAGGAGTG TTCTGGCCCCA GGGGATCTGT GAACAGGCTG GGAAGCATCT
 CON 1018 ACTGTTTTTC CTAAGGAGTG TTCTGGCCCCA GGGGATCTGT GAACAGGCTG GGAAGCATCT
 5 93 CAAGATCTTT CCAGGGTTAT ACTTACTAGC ACACAGCATG ATCATTACGG AGTGAATTAT
ES - 2F

CON 1078 CAAGATCTTT CCAGGGTTAT ACTTACTAGC ACACAGCATG ATCATTACGG AGTGAATTAT
 5 153 CTAATCAACA TCATCCTCAG TGTCCTTGCC CATACTGAAA TTCATTTCCTC ACTTTTGTGC
 CON 1138 CTAATCAACA TCATCCTCAG TGTCCTTGCC CATACTGAAA TTCATTTCCTC ACTTTTGTGC
 5 213 CCATTCTCAA GACCTCAAAA TGTCATTCCA TTAATATCAC AGGATTAACT TTTTTTTTTTA
 CON 1198 CCATTCTCAA GACCTCAAAA TGTCATTCCA TTAATATCAC AGGATTAACT TTTTTTTTTTA
 5 273 ACCTGGAAGA ATTCAATGTT ACATGCAGCT ATGGGAATTT AATTACATAT TTTGTTTTTCC
 CON 1258 ACCTGGAAGA ATTCAATGTT ACATGCAGCT ATGGGAATTT AATTACATAT TTTGTTTTTCC
 5 333 AGTGCAAAGA TGAATAAGTC CTTTATCCCT CCCCTTTGTT TGATTTTTTT TCCAGTATAA
 CON 1318 AGTGCAAAGA TGAATAAGTC CTTTATCCCT CCCCTTTGTT TGATTTTTTT TCCAGTATAA

FIG - 26

5 393 AGTTAAATG CTTAGCCTTG TACTGAGGCT GTATACAGCA CAGCCTCTCC CCATCCCTCC
CON 1378 AGTTAAATG CTTAGCCTTG TACTGAGGCT GTATACAGCA CAGCCTCTCC CCATCCCTCC

5 453 AGCCTTATCT GTCATCACCA TCAACCCCTC CCATNYSACC TAAACAAAAT CTAACCTGTA
CON 1438 AGCCTTATCT GTCATCACCA TCAACCCCTC CCATNYSACC TAAACAAAAT CTAACCTGTA

5 513 ATTCCTTGAA CATGTCAGGN CATACTTTC TCCTTCTGCC TGAGAAAGCTC TTCCTTGCTC
CON 1498 ATTCCTTGAA CATGTCAGGN CATACTTTC TCCTTCTGCC TGAGAAAGCTC TTCCTTGCTC

5 573 CTTAANTCTA GAATGATGTA AAGTTTGGAA TAAGTTGACT ATCTTACTTC ATGCAAAGAA
CON 1558 CTTAANTCTA GAATGATGTA AAGTTTGGAA TAAGTTGACT ATCTTACTTC ATGCAAAGAA

FILE - 2H

5 633 GGGACACATA TGAGATTTCAT CATCACATGA GACAGCAAAT ACTAAAAGTG TAATTTGATT
CON 1618 GGGACACATA TGAGATTTCAT CATCACATGA GACAGCAAAT ACTAAAAGTG TAATTTGATT

5 693 ATAAGAGTTT AGATAAATAT ATGAAATGCA AGAKCCACAG AGGGAATGTT TATGGGGCAC
CON 1678 ATAAGAGTTT AGATAAATAT ATGAAATGCA AGAKCCACAG AGGGAATGTT TATGGGGCAC

5 753 GTTTGTAAGC CTGGGATGTG AAGMAAAGGC AGGGAACCTC ATAGTATCTT ATATAATATA
CON 1738 GTTTGTAAGC CTGGGATGTG AAGMAAAGGC AGGGAACCTC ATAGTATCTT ATATAATATA

5 813 CTTCAATTTCT CTATCTCTAT CACAATATCC AACAAAGCTTT TCACAGAATT CATGCAGTGC
CON 1798 CTTCAATTTCT CTATCTCTAT CACAATATCC AACAAAGCTTT TCACAGAATT CATGCAGTGC

5 873 AAATCCCCAA AGGTAACCTT TATCCATTTC ATGGTGAGTG CGCTTTAGAA TTTTGGCAAA

Final - 21

CON 1858 AAATCCCCAA AGTAACCTT TATCCATTTC ATGGTGAGTG CGCTTTAGAA TTTTGGCAAA

5 933 TCATACTGGT CACTTATCTC AACTTTGAGA TGTGTTTGTC CTTGTAGTTA ATTGAAAGAA

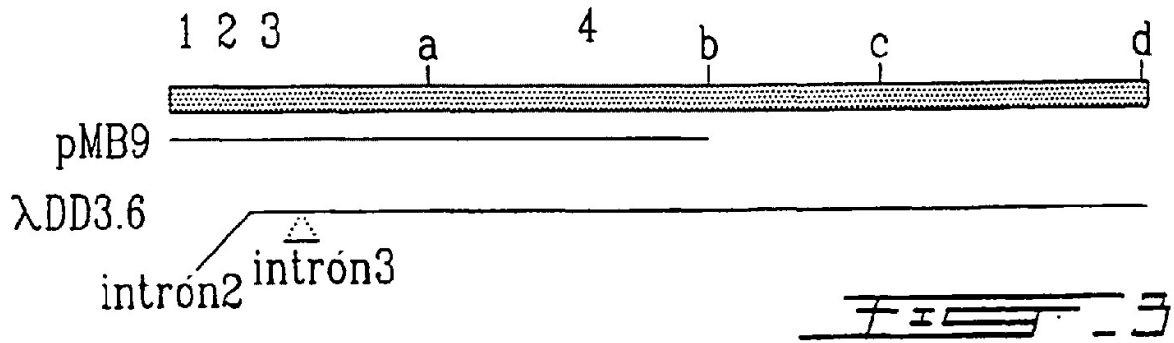
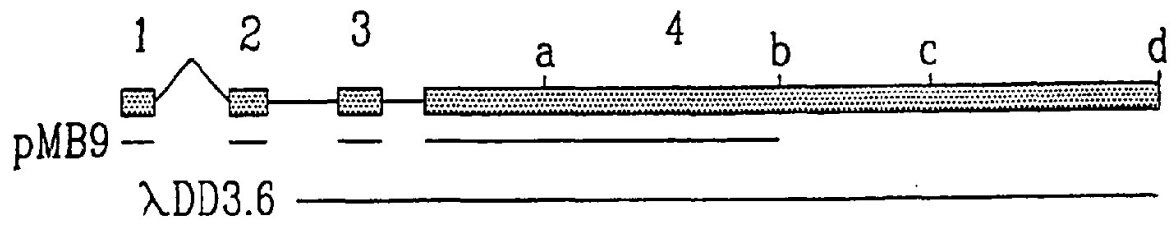
CON 1918 TCATACTGGT CACTTATCTC AACTTTGAGA TGTGTTTGTC CTTGTAGTTA ATTGAAAGAA

5 993 ATAGGGCACT CTTgtgagcc actttagggg tcaactcctgg caataaagaa tttacaaaga

CON 1978 ATAGGGCACT CTTgtgagcc actttagggg tcaactcctgg caataaagaa tttacaaaga

FILE - 2J

Clones de ADNc de PCA3



^{S1}
 PE | RACE 5' | RACE 5' ADNC
 ▼ ▼ ▼
 ...g ACAGAAG AAATAGCAAG TGCCGAGAAG CTGGCATCAG
 L→ TSS L→ pMB9
 AAAAACAGAG GGGAGATTG TGTGGCTGCA GCCGAGGGAG ACCAGGAAGA
 TCTGCATGGT GGAAGGACC TGATGATACA GAG gt ...
 sitio de ajuste

Fig. 4

```

286
285
121 469 1008 2067 2623
1 120 468 1007 2066 2622 3582
=====
>E1>E2>E3->-EXON-4A->-EXON-4B->-EXON-4C->-EXON-4D----->
>PCA3>
1|-----|3582
401 553
FE-5A

```

Secuencia de ADNc de PCA3 y proteína PCA3:

```

1 ACAGAAAGAAA TAGCAAGTGC CGAGAAGCTG GCATCAGAAA AACAGAGGGG AGATTGTGT
61 GGCTGCAGCC GAGGGAGACC AGGAAGATCT GCATGGTGGG AAGGACCTGA TGATACAGAG
121 GAATTACAAC ACATATACTT AGTGTTTCAA TGAACACCAA GATAAATAAG TGAAGAGCTA
181 GTCCGCTGTG AGTCTCCTCA GTGACACAGG GCTGGATCAC CATCGACGGC ACTTCTGAG
241 TACTCAGTGC AGCAAAGAAA GACTACAGAC ATCTCAATGG CAGGGGTGAG AAATAAGAAA
301 GGCTGCTGAC TTTACCATCT GAGGCCACAC ATCTGCTGAA ATGGAGATAA TTAACATCAC
FE-5B

```

361 TAGAAACAGC AAGATGACAA TATAATGTCT AAGTAGTGAC **ATGTTTTTGC** ACATTTCAG
 PCA3 1 M F L H I S S
 421 CCCCCTTTAA TATCCACACA CACAGGAAGC ACAAAGGAA GCACAGAGAT CCCTGGGAGA
 PCA3 8 P F K Y P H T Q E A Q K E A Q R S L G E
 481 AATGCCCCGC CGCCATCTTG GGTCAATCGAT GAGCCTCGCC CTGTGCCTGG TCCCCTTGT
 PCA3 28 M P G R H L G S S M S L A L C L V P L V
 541 GAGGGAAGGA CATTAGAAAA TGAATTGATG TGTTCCTTAA AGGATGGGCA GGAAAAACAGA
 PCA3 48 R E G H *
 601 TCCTGTTGTG GATATTTATT TGAACGGGAT TACAGATTG AAATGAAGTC ACAAAGTGAG
 661 CATTACCAAT GAGAGGAAAA CAGACGAGAA AATCTTGATG GCTTCACAAG ACATGCAACA
 721 AACAAAATGG AATACTGTGA TGACATGAGG CAGCCAAGCT GGGGAGGAGA TAACCACGGG
 781 GCAGAGGGTC AGGATTCTGG CCCTGCTGCC TAAACTGTGC GTTCATAACC AAATCATTC
 841 ATATTTCTAA CCCTCAAAAC AAAGCTGTTG TAATACTGA TCTCTACGGT TCCTTCTGGG
 901 CCCAACATTC TCCATATATC CAGCCACACT CATTTTAAAT ATTAGTTCC CAGATCTGTA
 961 CTGTGACCTT TCTACACTGT AGAATAACAT TACTCATTTT GTTCAAAGAC CCTTCGTGT
 1021 GCTGCCCTAAT ATGAGCTGA CTGTTTTTCC TAAGGAGTGT TCTGGCCCGG GGAATCTGTG

FIG - 5C

1081 AACAGGCTGG GAAGCATCTC AAGATCTTTC CAGGGTTATA CTTACTAGCA CACAGCATGA
 1141 TCATTACGGA GTGAATTATC TAATCAACAT CATCCTCAGT GTCTTTGCCC ATACTGAAAT
 1201 TCATTTCCCA CTTTGTGCC CATTCTCAAG ACCTCAAAAT GTCATTCCAT TAATATCACA
 1261 GGATTAACTT TTTTTTTAA CCTGGAAGAA TTCAAATGTTA CATGCAGCTA TGGGAATTTA
 1321 ATTACATATT TTGTTTTCCA GTGCAAAAGAT GACTAAGTCC TTTATCCCTC CCCTTTGTTC
 1381 GATTTTTTTT CCAGTATAAA GTTAAATGC TTAGCCTTGT ACTGAGGCTG TATACAGCAC
 1441 AGCCTCTCCC CATCCCTCCA GCCTTATCTG TCATCACCAT CAACCCCTCC CATAACACCT
 1501 AAACAAAATC TAACTTGTA TCCCTTGAAC ATGTCAGGAC ATACATTATT CCTTCTGCCT
 1561 GAGAAGCTCT TCCTTGTCTC TTAAATCTAG AATGATGTAA AGTTTGAAT AAGTTGACTA
 1621 TCCTTACTTCA TGCAAGAAG GGACACATAT GAGATTCATC ATCACATGAG ACAGCAAATA
 1681 CTAATAAGTGT AATTGATTA TAAGAGTTTA GATAAATATA TGAAATGCAA GAGCCACAGA
 1741 GGAATGTTT ATGGGACAG TTTGTAAGCC TGGGATGTA AGCAAAGGCA GGAACCTCA
 1801 TAGTATCTTA TATAATATAC TTCATTCTC TATCTCTATC ACAATATCCA ACAAGCTTTT
 1861 CACAGAAATC ATGCAGTGCA AATCCCCAAA GGTAACCTTT ATCCATTCA TGGTGAGTGC

755 - 50

1921 GCTTTAGAAT FTTGGCAAAT CATACTGGTC ACTTATCTCA ACTTTGAGAT GTGTTTGTCC
 1981 TTGTAGTTAA TTGAAAGAAA TAGGGCACTC TTGTGAGCCA CTTTAGGGTT CACTCCTGGC
 2041 AATAAAGAAT TTACAAAGAG CTACTCAGGA CCAGTTGTTA AGAGCTCTGT GTGTGTGTGT
 2101 GTGTGTGTGT GAGTGTACAT GCCAAAGTGT GCCTCTCTCT CTTGACCCAT TATTTCAGAC
 2161 TTAAACAAG CATGTTTCA AATGGCACTA TGAGCTGCCA ATGATGTATC ACCACCATAT
 2221 CTCATTATTC TCCAGTAAAT GTGATAATAA TGTCATCTGT TAACATAAAA AAAGTTTGAC
 2281 TTCACAAAAG CAGCTGAAA TGGACAACCA CAATATGCAT AAATCTAACT CCTACCATCA
 2341 GCTACACACT GCTTGACATA TATTGTTAGA AGCACCTCGC ATTTGTGGGT TCTCTTAAAGC
 2401 AAAATACTTG CATTAGGTCT CAGCTGGGGC TGTGCATCAG GCGGTTTGAG AAATATTCAA
 2461 TTCTCAGCAG AAGCCAGAAT TTGAATTCCC TCATCTTTTA GGAATCATTT ACCAGGTTTG
 2521 GAGAGGATTC AGACAGCTCA GTGCTTTCA CTAATGTCTC TGAACCTCTG TCCCCTCTTG
 2581 TGTTCAATGGA TAGTCCAATA AATAATGTTA TCTTTGAACT GATGCTCATA GGAGAGAATA
 2641 TAAGAAGTCT GAGTGATATC AACATFAGGG ATTCAAAGAA ATATTAGATT TAAGCTCACA
 2701 CTGGTCAAAA GGAACCAAGA TACAAAGAAC TCTGAGCTGT CATCGTCCCC ATCTCTGTGA

7151-5E

2761 GCCACAACCA ACAGCAGGAC CCAACGCATG TCTGAGATCC TTAATCAAG GAAACCAGTG
 2821 TCATGAGTTG AATTCTCCTA TTATGGATGC TAGCTTCTGG CCATCTCTGG CTCTCCTCTT
 2881 GACACATATT AGCTTCTAGC CTTTGCTTCC ACGACTTTTA TCTTTTCTCC AACACATCGC
 2941 TTACCAATCC TCTCTCTGCT CTGTGCTTT GGACTTCCCC ACAAGAATTT CAACGACTCT
 3001 CAAGTCTTTT CTTCCATCCC CACCACTAAC CTGAATTGCC TAGACCCCTTA TTTTATTAA
 3061 TTTCCAATAG ATGCTGCCCTA TGGGCTAATA TTGCTTTAGA TGAACATTAG ATATTAAAG
 3121 TCTAAGAGGT TCAAAATCCA ACTCATATC TTCCTTTCT TTCACCTCCC CTGCTCCTCT
 3181 CCCTATATTA CTGATTGACT GAACAGGATG GTCCCCAAGA TGCCAGTCAA ATGAGAAACC
 3241 CAGTGGCTCC TTGTGGATCA TGCATGCAAG ACTGCTGAAG CCAGAGGATG ACTGATTACG
 3301 CCTCATGGGT GGAGGGGACC ACTCCTGGGC CTTCGTGATT GTCAGGAGCA AGACCTGAGA
 3361 TGCTCCCTGC CTTCAGTGTC CTCGCACTC CCCCTTTCTA ATGAAGATCC ATAGAAATTG
 3421 CTACATTTGA GAATTCCAAT TAGGAACTCA CATGTTTAT CTGCCCTATC AATTTTTAA
 3481 ACTTGCTGAA AATTAAGTTT TTTCAAAATC TGTCTTGTA AATTACTTTT TCTTACAGTG
 3541 TCTTGGCATA CTATATCAAC TTTGATTCTT TGTTACAACT TT

ES - 5F