



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 324 503**

51 Int. Cl.:

C12N 15/00 (2006.01)	C12N 15/12 (2006.01)
C12N 5/12 (2006.01)	C12N 1/21 (2006.01)
C07K 14/47 (2006.01)	C07K 16/18 (2006.01)
C12Q 1/68 (2006.01)	G01N 33/577 (2006.01)
A61K 39/395 (2006.01)	A01K 67/027 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **98916696 .2**

96 Fecha de presentación : **09.04.1998**

97 Número de publicación de la solicitud: **1007650**

97 Fecha de publicación de la solicitud: **14.06.2000**

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

30 Prioridad: **10.04.1997 US 41836 P**

45 Fecha de publicación de la mención BOPI:
07.08.2009

45 Fecha de la publicación del folleto de la patente:
07.08.2009

73 Titular/es: **Stichting Katholieke Universiteit
University Medical Centre Nijmegen
Geert Grooteplein 10
6525 GA Nijmegen, NL
The John Hopkins University**

72 Inventor/es: **Bussemakers, Marion, J., G. y
Isaacs, William Brewster**

74 Agente: **Isern Jara, Jorge**

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 **Ámbito de la invención**

La presente invención se refiere, en general, a un antígeno de cáncer de próstata, el PCA3. En particular, la presente invención se refiere a moléculas de ácido nucleico que codifican a la proteína del PCA3; a proteínas y polipéptidos de PCA3 purificados; a moléculas de ácido nucleico recombinante; a células que contienen moléculas de ácido nucleico recombinante; a anticuerpos que tienen afinidad de fijación específica sobre proteínas y polipéptidos de PCA3; a hibridomas que contienen los anticuerpos; sondas de ácido nucleico para la detección de ácidos nucleicos que codifican a las proteínas PCA3; un método para detectar ácidos nucleicos que codifican a proteínas o polipéptidos de PCA3 de una muestra; kits que contienen sondas de ácido nucleico o anticuerpos; bioensayos que utilizan la secuencia de ácidos nucleicos, la proteína o los anticuerpos de esta invención para diagnosticar, evaluar o pronosticar a un mamífero que sufre cáncer de próstata; usos terapéuticos; y métodos para prevenir el cáncer de próstata en un animal.

Antecedentes de la invención

El cáncer de próstata es la enfermedad maligna de diagnóstico más frecuente y la principal causa de mortalidad relacionada con el cáncer en la población masculina del mundo occidental. Cuando este carcinoma se propaga de forma local o a una posición alejada del origen, no se dispone de terapia curativa. Por consiguiente, los esfuerzos para controlar la enfermedad (es decir, para disminuir la mortalidad por cáncer de próstata) se han centrado en aumentar la detección del cáncer cuando todavía está localmente confinado y es potencialmente curable. Los estudios dirigidos a la detección precoz del cáncer de próstata han puesto de manifiesto un aumento apreciable de la detección de los cánceres de próstata confinados en el órgano y potencialmente curables. Sin embargo, todavía no se ha demostrado que un mayor índice de detección se traduzca en una disminución del índice de mortalidad específica por cáncer de próstata. Por otro lado, no hay indicios de que el diagnóstico precoz permita disminuir el índice de mortalidad. Tanto en Estados Unidos como en Europa se hallan todavía vivos los debates sobre la eficacia y la aceptabilidad de los programas de exploración, el tema del diagnóstico excesivo y del tratamiento excesivo y las opciones que tiene un tratamiento precoz de reducir la morbilidad y la mortalidad por cáncer de próstata, por lo cual el tema de la detección precoz del cáncer de próstata se ha convertido en objeto de controversia (Schröder, Urology 46, 6270, 1995).

Las mediciones de las concentraciones de enzimas marcadoras prostáticas en suero tienen un valor reconocido para la detección, diagnosis y gestión clínica del cáncer de próstata. Las dos enzimas marcadoras prostáticas empleadas con mayor frecuencia son la fosfatasa ácida prostática (PAP) y antígeno prostático específico (PSA). Normalmente, las dos enzimas son secretadas por las células epiteliales de la próstata al líquido seminal, pero en pacientes que sufren la enfermedad prostática se vierten en el torrente circulatorio, donde pueden detectarse mediante ensayos inmunológicos (Armbruster, Clin. Che. 39, 181-95, 1993).

La fosfatasa ácida prostática, uno de los primeros marcadores séricos de la próstata, tiene una función todavía no determinada y es uno de los componentes proteicos más abundantes en las secreciones prostáticas humanas. El uso de la PAP como marcador de tumores prostáticos es complicado debido a las similitudes estructurales ya publicadas entre la fosfatasa ácida específica de la próstata y la fosfatasa ácida lisosómica presente en todos los tejidos. Además, hay una tendencia a niveles más bajos de PAP-mRNA y de proteína en el cáncer de próstata si se compara con la hiperplasia benigna de próstata (BPH). En los últimos años, en la gestión clínica del cáncer de próstata, las mediciones de la PAP se están reemplazando por mediciones del PSA en suero.

El antígeno prostático específico ha sido identificado por varios grupos de científicos en la década de los años 1970 como proteína específica de la próstata que se halla en el líquido seminal. En 1979 se purificó como antígeno existente en el tejido de un cáncer de próstata. Las investigaciones posteriores han puesto de manifiesto que el PSA se produce exclusivamente en las células epiteliales columnares de la próstata y de las glándulas periuretrales. Sin embargo, el epitelio prostático normal y el tejido hiperplásico benigno producen más PSA-mRNA y proteína que el tejido del cáncer de próstata. Se ha constatado además que la pérdida de la diferenciación de los carcinomas prostáticos está asociada con una disminución del nivel del PSA intraprostático.

Las anomalías de la arquitectura prostática que surgen como resultado de la enfermedad prostática conducen a un mayor vertido de PSA (y de PAP) al suero y las mediciones del PSA en el suero se convierten en marcadores del cáncer de próstata. A pesar de que los estudios tempranos indicaron que el ensayo de diagnóstico del PSA podría estar obstaculizado por el hecho de que carece de especificidad y no distingue entre BPH y cáncer de próstata, el ensayo del PSA se implantó en 1986 y revolucionó la gestión de los pacientes que sufren cáncer de próstata. El mejor conocimiento de la especificidad de órgano del PSA y la relación entre los niveles elevados de PSA en suero con la enfermedad de la próstata así como la mejora de las técnicas de biopsia y de evaluación histológica han conducido a una mejor apreciación del valor clínico de los ensayos del PSA, una utilidad que todavía no ha podido ser igualada por ningún otro marcador tumoral (prostático).

La clonación del gen que codifica al PSA ha puesto de manifiesto que es un componente del grupo de genes de calicreína humana y se ha traducido en el desarrollo de nuevas estrategias para el uso del PSA como marcador: se emplea el método muy sensible de la reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR) para detectar números extremadamente pequeños de células prostáticas malignas en muestras de sangre de pacientes que sufren cáncer de próstata y puede proporcionar una herramienta sensible para identificar a pacientes que tienen una enfermedad micrometastásica (Moreno y col., Cancer Res. 52, 6110-12, 1992; y Katz y col., Urology 43, 765-75, 1994).

El antígeno de membrana específico de la próstata (PSM) se identificó inicialmente empleando un anticuerpo desarrollado para inmunizar ratones con la fracción de membrana de células de adenocarcinoma prostático humano LNCaP. Al igual que la PAP y el PSA, el PSM puede detectarse en la próstata normal, en la BPH y en el cáncer de próstata y está ausente en la mayoría de tejidos restantes. También para el PSM se han desarrollado estudios de RT-PCR para detectar células circulantes de cáncer de próstata, sin embargo, se requieren más investigaciones para establecer la utilidad del PSM como marcador del cáncer prostático.

Resumiendo, el PSA se reconoce actualmente como el principal marcador del cáncer de próstata y es útil para explorar las poblaciones seleccionadas de pacientes que presentan síntomas que indican cáncer de próstata y para el seguimiento de los pacientes después de la terapia, en especial después de la prostatectomía quirúrgica (los niveles medibles de PSA indican una enfermedad residual o una metástasis y el aumento de las concentraciones de PSA indican la recaída en la enfermedad). Los puntos débiles significativos del PSA como marcador tumoral son que (1) el PSA no siempre es capaz de distinguir entre cáncer de próstata y la BPH; y (2) que su expresión disminuye con la pérdida de diferenciación de los carcinomas.

Teniendo en cuenta que el cáncer de próstata avanzado continúa siendo una enfermedad letal que alcanza a una proporción muy significativa de la población masculina, sigue habiendo necesidad de desarrollar nuevas modalidades de tratamiento y de diagnóstico para (la última fase del) cáncer de próstata.

La presente invención intenta satisfacer esta y otras demandas.

La presente descripción menciona un gran número de documentos, cuyos contenidos se incorporan a la presente como referencias.

Resumen de la invención

La invención proporciona, en general, moléculas de ácido nucleico aisladas que codifican al PCA3 o a fragmentos del mismo.

En particular, la presente invención proporciona los elementos siguientes:

(i) Una molécula de ácido nucleico asociada con el cáncer de próstata y derivada de secuencias de exón de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), que consta de una secuencia de polinucleótido elegida entre el grupo formado por:

(a) una secuencia de nucleótidos del exón 1 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 1-98 de la SEQ ID NO: 1;

(b) una secuencia de nucleótidos del exón 2 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 99-263 de la SEQ ID NO: 1;

(c) una secuencia de nucleótidos del exón 3 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 264-446 de la SEQ ID NO: 1;

(d) una secuencia de nucleótidos del exón 4a de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 447-985 de la SEQ ID NO: 1;

(e) una secuencia de nucleótidos del exón 4b de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 986-2037 de la SEQ ID NO: 1;

(f) una secuencia de nucleótidos que contiene secuencias de combinaciones de exones de las secuencias de nucleótidos de (a), (b), (c), (d) y (e);

(g) una secuencia de nucleótidos del exón 1 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 1-120 de la SEQ ID NO: 6;

(h) una secuencia de nucleótidos del exón 2 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 121-285 de la SEQ ID NO: 6;

ES 2 324 503 T3

- (i) una secuencia de nucleótidos del exón 3 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 286-468 de la SEQ ID NO: 6;
- 5 (j) una secuencia de nucleótidos del exón 4a de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 469-1007 de la SEQ ID NO: 6;
- (k) una secuencia de nucleótidos del exón 4b de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 1008-2066 de la SEQ ID NO: 6;
- 10 (l) una secuencia de nucleótidos del exón 4c de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 2067-2622 de la SEQ ID NO: 6;
- (m) una secuencia de nucleótidos del exón 4d de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 2623-3582 de la SEQ ID NO: 6;
- 15 (n) una secuencia de nucleótidos que contiene secuencias de combinaciones de exones de las secuencias de nucleótidos de (g), (h), (i), (j), (k), (l) y (m);
- (o) una secuencia de nucleótidos que es completamente complementaria de cualquiera de las secuencias de nucleótidos de (a) a (n); y
- 20 (p) una secuencia de nucleótidos que se hibrida específicamente en condiciones de gran restricción a 68°C de cualquiera de las secuencias de nucleótidos de (a) a (o), dicha secuencia de nucleótidos contiene de 10 a 1000 nucleótidos.
- 25 (ii) La molécula de ácido nucleico del elemento (i), en la que dicha secuencia de nucleótidos de (p) consta de:
- (a) de 10 a 50 nucleótidos;
- (b) de 10 a 35 nucleótidos; o
- 30 (c) de 20 a 35 nucleótidos.
- (iii) La molécula de ácido nucleico del elemento (i), en la que dicha secuencia de nucleótidos de (p) se hibrida con una secuencia definida en la SEQ ID NO: 1.
- 35 (iv) La molécula de ácido nucleico del elemento (i), en la que dicha secuencia de nucleótidos de (f) o (n) se hibrida con una secuencia definida en la SEQ ID NO: 6.
- (v) Un método para detectar una molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3) en una muestra que consiste en:
- (a) poner en contacto dicha muestra con la molécula de ácido nucleico según uno cualquiera de los elementos (i)-(iv); y
- 45 (b) detectar la presencia de dicha molécula fijada sobre la molécula de ácido nucleico de PCA3, empleando una secuencia de nucleótidos derivada de dicha molécula de ácido nucleico del antígeno de cáncer de próstata (PCA3).
- (vi) El método del elemento (v), en el que dicha molécula de ácido nucleico de PCA3 contiene los nucleótidos 1-98 de la SEQ ID NO: 1 o los nucleótidos 1-120 de la SEQ ID NO: 6.
- 50 (vii) El método del elemento (v) o (vi), en el que dicha molécula de ácido nucleico de PCA3 contiene los nucleótidos 264-446 de la SEQ ID NO: 1 o los nucleótidos 286-468 de la SEQ ID NO: 6.
- 55 (viii) El método del elemento (v), (vi) o (vii), en el que dicha secuencia de (b) está presente en la SEQ ID NO: 4.
- (ix) El método de uno cualquiera de los elementos (v)-(viii), en el que dicha detección consiste en una detección de un radiomarcador, de fluorescencia, de quimioluminiscencia o cualquier combinación de las mismas.
- 60 (x) Una combinación de cebadores para la amplificación de una secuencia de exón de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), en la que dichos cebadores se hibridan específicamente con la molécula de ácido nucleico del elemento (i).
- 65 (xi) La combinación de cebadores del elemento (x), en la que cada uno de dichos de cebadores está formado por 10-50 nucleótidos.

ES 2 324 503 T3

(xii) El método según uno cualquiera de los elementos (v)-(ix) que consiste en:

(a) amplificar dicha molécula de ácido nucleico de PCA3 en dicha muestra, poner en contacto dicha muestra con dicha combinación de cebadores según el elemento (x) o (xi); y

(b) detectar la presencia de dicha molécula de ácido nucleico de PCA3 visualizando el producto de la amplificación de (a).

(xiii) El método del elemento (xii), en el que dicha muestra se elige entre el grupo formado por:

(a) una muestra que contiene tejido o células de próstata;

(b) el RNA de tejido o células humanas;

(c) el DNA de tejido o células humanas; y

(d) orina, sangre o suero humanos.

(xiv) El método del elemento (xiii), en el que dicha muestra es orina humana.

(xv) Una molécula de ácido nucleico recombinante que contiene la secuencia de un vector y la molécula de ácido nucleico según uno cualquiera de los elementos (i)-(iv).

(xvi) La molécula de ácido nucleico recombinante del elemento (xv), que consta de un promotor eficaz para iniciar la transcripción de dicha molécula de ácido nucleico en una célula hospedante.

(xvii) Una célula que contiene la molécula de ácido nucleico de uno cualquiera de los elementos (i)-(iv), o la molécula de ácido nucleico recombinante del elemento (xv) o (xvi).

(xviii) Un organismo no humano que contiene la molécula de ácido nucleico recombinante según el elemento (xv) o (xvi).

(xix) La combinación de cebadores del elemento (x), en la que dichos cebadores se hibridan con secuencias de combinaciones de exones de dicha molécula de ácido nucleico del antígeno de cáncer de próstata (PCA3).

(xx) La molécula de ácido nucleico de uno cualquiera de los elementos (i)-(iv), el método de uno cualquiera de los elementos (v)-(ix), (xii), (xiii) y (xiv), o la combinación de cebadores del elemento (x), (xi) o (xix), en los que dicha molécula de ácido nucleico o cebadores están formados por lo menos por 15 nucleótidos.

(xxi) La molécula de ácido nucleico, el método o la combinación de cebadores del elemento (xx), en los que dicha molécula de ácido nucleico o dichos cebadores tienen un tamaño elegido entre el grupo formado por:

(a) 20-35 nucleótidos;

(b) 20-50 nucleótidos; y

(c) 20-100 nucleótidos.

(xxii) La molécula de ácido nucleico, el método o la combinación de cebadores del elemento (xxi), en los que la molécula de ácido nucleico o dichos cebadores están formados por 20-35 nucleótidos.

La presente invención se define en las reivindicaciones.

Se describen también aquí polipéptidos purificados que codifican al PCA3 o a una porción del mismo que se fija sobre un epítipo.

Se describe también aquí un kit para detectar la presencia del ácido nucleico que codifica al PCA3 en una muestra. Se describe también aquí un kit de diagnóstico formado por un primer recipiente que contiene el anticuerpo antes descrito y un segundo recipiente que contiene un conjugado formado por un conjugado de un reactivo de fijación del anticuerpo monoclonal y un marcador.

Se describe también aquí una molécula antisentido del ácido nucleico de PCA3.

Se describe también aquí un anticuerpo que tiene una afinidad de fijación específica con el PCA3 o una porción del mismo que lleva un epítipo. Se describe también aquí un método para detectar anticuerpos que tengan afinidad de fijación específica sobre el PCA3.

Se describen también aquí reactivos inmunogénicos para inducir la protección contra las células de cáncer de próstata que expresan al PCA3. Con preferencia, tales reactivos inmunogénicos son polipéptidos que codifican al PCA3, una porción antigénica de los mismos, proteínas de fusión que codifican al PCA3 o proteínas de fusión que codifican a las porciones antigénicas del PCA3. En tal forma de ejecución, estos reactivos inmunogénicos actuarían

como vacunas.

Se describe también aquí un hibridoma que produce al anticuerpo monoclonal antes descrito.

Se describen también aquí métodos para el uso terapéutico, que impliquen la totalidad o una parte de (1) una secuencia de ácido nucleico que codifique al PCA3, (2) moléculas de ácido nucleico de PCA3 antisentido, (2) proteína de PCA3 o (4) anticuerpos de PCA3.

Otros objetos y ventajas de la presente invención resultarán manifiestos después de la descripción que sigue.

Definiciones

En la descripción que sigue, se emplea con frecuencia un gran número de términos de la tecnología del DNA (rDNA) recombinante. Con el fin de facilitar una comprensión clara y consistente de la descripción y reivindicaciones, incluido el alcance que se pretende dar a los términos, se establecen las definiciones siguientes.

Molécula aislada de ácido nucleico. Una "molécula aislada de ácido nucleico", tal como se entiende y emplea en general, indica un polímero de nucleótidos e incluye, pero no se limita a: DNA y RNA. La molécula "aislada" de ácido nucleico se purifica a partir de su estado natural "*in vivo*".

DNA recombinante. Cualquier molécula de DNA formada por unión de segmentos de DNA de diferentes fuentes y obtenida aplicando una tecnología de DNA recombinante (también conocida como ingeniería genética molecular).

Segmento de DNA. Un segmento de DNA, tal como entiende y emplea en general, indica una molécula que contiene un tramo lineal de nucleótidos, dichos nucleótidos están presentes en una secuencia que puede codificar, a través del código genético, a una molécula que contiene una secuencia lineal de restos aminoácido, que se denomina proteína, fragmento de proteína o polipéptido.

Gen. Una secuencia de DNA relativa a una cadena de polipéptido o proteína individual que, tal como se emplea aquí, incluye a los extremos 5' y 3' sin traducir. El polipéptido puede codificarse con una secuencia de longitud completa o con cualquier porción de una secuencia codificante, en el supuesto de que se conserve la actividad funcional de la proteína.

DNA (cDNA) complementario. Las moléculas recombinantes de ácido nucleico sintetizadas por transcripción inversa del RNA mensajero ("mRNA").

Gen estructural. Una secuencia de DNA que se transcribe al mRNA que después se traduce en una secuencia de aminoácidos característica de un polipéptido específico.

Endonucleasa de restricción. Una endonucleasa de restricción (también llamada enzima de restricción) es una enzima que tiene la capacidad de reconocer una secuencia específica de bases (normalmente de una longitud de 4, 5 ó 6 pares de bases) dentro de una molécula de DNA y de romper la molécula de DNA por cualquier lugar, en el que aparezca esta secuencia. Por ejemplo, el EcoRI reconoce la secuencia de bases GAATTC/CTTAAG.

Fragmento de restricción. Las moléculas DNA producidas por digestión con una endonucleasa de restricción se denominan fragmentos de restricción. Cualquier genoma puede digerirse con una endonucleasa de restricción concreta para convertirse en un conjunto discreto de fragmentos de restricción.

Electroforesis a través de gel de agarosa. Para detectar un polimorfismo en la longitud de los fragmentos de restricción se requiere un método analítico que fraccione las moléculas de DNA de doble hebra con arreglo al tamaño. La técnica aplicada con mayor frecuencia (aunque no es la única) para conseguir tal fraccionamiento es la electroforesis a través de gel de agarosa. El principio, en el que se basa este método, es que las moléculas de DNA migran a través del gel como si fuera un tamiz, que retrasa el movimiento de las moléculas más voluminosas en mayor grado y el movimiento de las moléculas más pequeñas en un menor grado. Cabe señalar que cuanto menor es el fragmento de DNA, tanto mayor será su movilidad en la electroforesis a través del gel de agarosa.

Los fragmentos de DNA fraccionado por electroforesis a través de gel de agarosa pueden visualizarse directamente por un procedimiento de tinción, si el número de fragmentos incluidos en la muestra es pequeño. Los fragmentos de DNA de genomas pueden visualizarse con éxito. Sin embargo, la mayoría de genomas, incluido el genoma humano, contienen demasiadas secuencias de DNA para poder producir un modelo simple de fragmentos de restricción. Por ejemplo, el genoma humano se digiere con el EcoRI dando lugar aproximadamente a 1.000.000 fragmentos de DNA diferentes. Con el fin de visualizar un subconjunto pequeño de estos fragmentos puede aplicarse una metodología llamada procedimiento de hibridación Southern.

Procedimiento de transferencia Southern. La finalidad del procedimiento de transferencia Southern (también llamado transferencia Southern) es transferir físicamente el DNA fraccionado por electroforesis a través de gel de agarosa a un papel de filtro de nitrocelulosa u otra superficie o método apropiados, al tiempo que se conservan las posiciones relativas de los fragmentos de DNA obtenidos por el procedimiento de fraccionamiento. La metodología aplicada para conseguir la transferencia del gel de agarosa a la nitrocelulosa implica sacar al DNA del gel y pasarlo al papel de nitrocelulosa por capilaridad.

Hibridación de ácido nucleico. La hibridación del ácido nucleico depende del principio que dice que dos moléculas de ácido nucleico de hebra simple, que tienen secuencias de bases complementarias, volverán a formar una estructura termodinámicamente más favorable de doble hebra si se mezclan en condiciones apropiadas. La estructura de doble hebra se formará entre los dos ácidos nucleicos de hebra simple complementaria incluso si uno de ellos está inmovilizado sobre un filtro de nitrocelulosa. En el procedimiento de hibridación Southern surge esta última situación. Tal como se ha dicho anteriormente, el DNA del individuo que se quiere analizar se digiere con una endonucleasa de restricción, se fracciona por electroforesis a través de gel de agarosa, se convierte en la forma de hebra simple y se transfiere al papel de nitrocelulosa, dejándolo disponible para la nueva fusión con la sonda de hibridación. Los ejemplos de condiciones de hibridación podrán encontrarse en Ausubel, F.M. y col., Current protocols in Molecular Biology, John Wiley & Sons, Inc., Nueva York, NY (1989). Se incuba el filtro de nitrocelulosa a 68°C durante una noche con una sonda marcada en una solución que contienen formamida del 50%, alto contenido de sal (ya sea 5X SSC [20X: 3M NaCl/0,3M citrato trisódico] o 5X SSPE [20X: 3,6M NaCl/0,2M NaH₂PO₄/0,02M EDTA, pH 7,7]), 5X solución de Denhardt, 1% SDS y 100 µg/ml de DNA de esperma de salmón desnaturalizado. Siguen varios lavados en 0,2X SSC/0,1% SDS a una temperatura elegida en función de la restricción deseada: temperatura ambiente (restricción baja), 42°C (restricción moderada) o 68°C (restricción elevada). La temperatura elegida se determina en base a la temperatura de fusión (T_m) del híbrido de DNA.

Sonda de hibridación. Para visualizar una secuencia de DNA concreta en el procedimiento de hibridación Southern se hace reaccionar una molécula de DNA marcada o una sonda de hibridación con el DNA fraccionado fijado sobre el filtro de nitrocelulosa. Las zonas del filtro que llevan las secuencias de DNA complementarias de la sonda del DNA marcado se convierten también en marcadas como consecuencia de la reacción de nueva fusión. Se visualizan las zonas del filtro que presentan tal marcado. La sonda de hibridación se produce normalmente por clonación molecular de una secuencia específica de DNA.

Oligonucleótido u oligómero. Es una molécula formada por dos o más desoxirribonucleótidos o ribonucleótidos, con preferencia más de tres. Su tamaño exacto dependerá de muchos factores, que a su vez dependen de la función o uso final del oligonucleótido. Un oligonucleótido puede derivarse por métodos sintéticos o por clonación.

Amplificación de secuencia. Es un método para generar grandes cantidades de una secuencia deseada (diana). En general se fusionan uno o más cebadores de amplificación con una secuencia de ácido nucleico. Empleando enzimas apropiadas se amplifican las secuencias situadas adyacentes o entre los cebadores.

Cebador de amplificación. Un oligonucleótido que es capaz de fusionarse de modo adyacente a una secuencia diana y que sirve como punto de inicio para la síntesis de DNA, cuando se somete a condiciones en las que se inicia la síntesis de un producto de extensión de cebador, que es complementario de una hebra de ácido nucleico.

Molécula antisentido de ácido nucleico. Una "molécula antisentido de ácido nucleico" indica una molécula capaz de formar un dúplex o un tríplex estable con una porción de su secuencia de ácido nucleico diana (DNA o RNA). En la técnica ya se conocen el uso de moléculas antisentido de ácido nucleico y el diseño o modificación de tales moléculas, que se han descrito por ejemplo en WO 96/32966, WO 96/11266, WO 94/15646, WO 93/08845 y US-5,593,974. Las moléculas antisentido de ácido nucleico aquí descritas pueden derivarse de las secuencias de ácido nucleico de la presente invención y modificarse con arreglo a métodos bien conocidos. Por ejemplo, se pueden diseñar algunas moléculas antisentido para que sean más resistente a la degradación, para aumentar su afinidad con su secuencia diana, para efectuar su transporte hacia los tipos de células elegidos o los compartimentos celulares elegidos y/o para mejorar su solubilidad en lípidos empleando análogos de nucleótido y/o sustituyendo los fragmentos químicos elegidos de las mismas del modo ya conocido en general en la técnica.

Vector. Un DNA plásmido o fago o otra secuencia de DNA en la que puede insertarse el DNA para clonarse. El vector puede replicarse de forma autónoma en una célula hospedante y después puede caracterizarse con uno o un pequeño número de sitios de reconocimiento de endonucleasa, en los que dichas secuencias de DNA pueden cortarse de un modo determinable y en las que puede insertarse el DNA. El vector puede contener además un marcador idóneo para el uso en la identificación de células transformadas con el vector. Los marcadores son, por ejemplo, la resistencia a la tetraciclina o la resistencia a la ampicilina. La expresión "vehículo de clonación" se emplea algunas veces como sinónimo de "vector".

Expresión. La expresión es un proceso en el que un gen estructural produce un polipéptido. Implica la transcripción del gen en el mRNA y la traducción de tal mRNA en el o en los polipéptido(s).

Vector de expresión. Un vector o un vehículo similar para un vector de clonación, pero que es capaz de expresar un gen que se ha clonado en él, después de la transformación en hospedante. El gen clonado se coloca normalmente bajo el control de (es decir, se une operativamente a) ciertas secuencias de control, por ejemplo las secuencias de promotor.

ES 2 324 503 T3

Las secuencias de control de expresión variarán en función de si el vector se ha diseñado para expresar el gen unido operativamente en un hospedante procariota o eucariota y si puede contener además elementos de transcripción, tales como elementos de amplificador, secuencias de terminación, elementos de especificidad de tejido y/o sitios de inicio y de terminación de la traducción.

Derivado funcional. Un “derivado funcional” de una secuencia, ya sea proteína, ya sea ácido nucleico, es una molécula que posee una actividad biológica (funcional o estructural), que es sustancialmente similar a la actividad biológica de la proteína o de la secuencia de ácido nucleico. Un derivado funcional de una proteína puede contener modificaciones posteriores a la traducción, por ejemplo hidratos de carbono unidos mediante enlace covalente, en función de la necesidad de tales modificaciones para la realización de una función específica. El término “derivado funcional” incluye también a “fragmentos”, “segmentos”, “variantes”, “análogos” o “derivativos químicos” de una molécula.

Tal como se emplea aquí, una molécula se dice que es un “derivado químico” de otra molécula si contiene restos químicos adicionales, que normalmente no forman parte de la otra molécula. Tales restos pueden mejorar la solubilidad de la molécula, la absorción, la vida media biológica y similares. Los restos pueden también disminuir la toxicidad de la molécula, eliminar o atenuar cualquier efecto secundario molesto de la molécula y similares. Los restos capaces de provocar tales efectos se han descrito en el manual Remington’s Pharmaceutical Sciences (1980). Los procedimientos para insertar tales restos en una molécula ya son conocidos por los expertos en la materia.

Variante. Una “variante” de una proteína o ácido nucleico se emplea para indicar una molécula sustancialmente similar en estructura y actividad biológica a la proteína o al ácido nucleico. Por lo tanto, si se da la condición de que dos moléculas posean una actividad y puedan intercambiarse entre sí, entonces se considerarán variantes en el sentido del término aquí empleado, a pesar de que la composición o la estructura secundaria, terciaria o cuaternaria de una de las moléculas no sea idéntica a la existente en la otra, o de que la secuencia de aminoácidos o de nucleótidos no sea idéntica.

Alelo. Un “alelo” es una forma alternativa de un gen, que ocupa un lugar determinado dentro del cromosoma.

Mutación. Una “mutación” es cualquier cambio detectable del material genético, que se transmite a las células hijas y posiblemente incluso a las generaciones futuras, dando lugar a células mutantes o a individuos mutantes. Si los descendientes de una célula mutante dan lugar solamente a células somáticas de organismos multicelulares, entonces surge un puñado o una zona de células mutantes. Las mutaciones de la línea germinal de organismos de reproducción sexual pueden transmitirse con los gametos a la siguiente generación, dando lugar a un individuo que tendrá la nueva condición mutante tanto en las células somáticas como en las germinales. Una mutación puede ser cualquier cambio no natural detectable (o una combinación de los mismos), que afecte a la constitución química o física, a la mutabilidad, replicación, función fenotípica o recombinación de uno o más desoxirribonucleótidos; los nucleótidos pueden añadirse, suprimirse, sustituirse, invertirse o trasponerse a nuevas posiciones con o sin inversión. Las mutaciones pueden surgir de forma espontánea o pueden inducirse experimentalmente con la aplicación de mutágenos. Una variación mutante de una molécula de ácido nucleico da lugar a una mutación. Un polipéptido mutante puede ser el resultado de una molécula de ácido nucleico mutante.

Especie. Una “especie” es un grupo de poblaciones naturales cultivadas juntas, de modo actual o potencial. Una variación de especie dentro de una molécula de ácido nucleico o proteína es un cambio en el ácido nucleico o en la secuencia de aminoácidos que tiene lugar entre las especies y que puede determinarse por secuenciación del DNA de la molécula en cuestión.

Electroforesis a través de un gel de poliacrilamida (PAGE). La técnica más empleada (aunque no es la única) para conseguir el fraccionamiento de polipéptidos por tamaños es la electroforesis a través de un gel de poliacrilamida. El principio en el que se basa este método es que las moléculas de polipéptido migran a través del gel como si fuera un tamiz que retarda en mayor grado el movimiento de las moléculas más voluminosas y en menor grado el movimiento de las moléculas más pequeñas. Cabe señalar que cuanto menor es el fragmento de polipéptido, tanto mayor será su movilidad en la electroforesis a través del gel de poliacrilamida. Tanto antes como durante la electroforesis, los polipéptidos están expuestos de modo continuo al detergente dodecil-sulfato sódico (SDS), en estas condiciones los polipéptidos se desnaturalizan. Los geles nativos se eluyen en ausencia de SDS. Los polipéptidos fraccionados por electroforesis a través de un gel de poliacrilamida pueden visualizarse directamente por un procedimiento de tinción, si el número de componentes polipéptido es pequeño.

Procedimiento de transferencia Western. La finalidad del procedimiento de transferencia Western (también llamado transferencia Western) es la de transferir los polipéptidos fraccionados por electroforesis a través de un gel de poliacrilamida a un papel de filtro de nitrocelulosa u otra superficie o método apropiados, al tiempo que se conservan las posiciones relativas de los polipéptidos resultantes del procedimiento de fraccionamiento. La transferencia se verifica seguidamente con un anticuerpo que se une específicamente al polipéptido de interés.

Purificado. Una proteína o un ácido nucleico “purificado” es una proteína o ácido nucleico que se ha separado de un componente celular. Proteínas o ácidos nucleicos “purificados” son los que se han purificado hasta un nivel de pureza que no existe en la naturaleza.

Sustancialmente puro. Una proteína o ácido nucleico “sustancialmente puro” es una preparación de proteína o ácido nucleico que carece de todos los demás componentes celulares.

Breve descripción de las figuras

Después de haber descrito la invención de forma general se acompaña ahora la referencia a las figuras anexas, que presentan a título ilustrativo una forma de ejecución preferida de la misma, en la que:

La figura 1 presenta la estructura genómica del gen PCA3.

La figura 2A-B presenta: (A) una estructura del cDNA del PCA3; (B) una secuencia de nucleótidos y aminoácidos de cDNA del PCA3 (SEQ ID NO: 1 y 2).

La figura 3A-B es una representación esquemática en la que comparan los clones pMB9 y λ DD3.6 de cDNA.

La figura 4 representa la situación de un sitio de inicio de transcripción (TSS) de PCA3. El sitio de inicio de transcripción se determina con los ensayos de extensión del cebador (PE), el mapeo de la S1-nucleasa (S1) y la amplificación rápida de los extremos 5' de cDNA (RACE).

La figura 5A-B representa: (A) una estructura de cDNA de PCA3; (B) una secuencia de nucleótidos y aminoácidos de cDNA del PCA3 (SEQ ID NO: 6 y 7); se han subrayado las señales de poli-adenilación putativa.

Otros objetos, ventajas y características de la presente invención resultarán evidentes con la lectura de la de siguiente descripción no restrictiva de las formas de ejecución preferidas, con referencia a las figuras anexas, que son ilustrativas y en modo alguno deberían tomarse como una limitación para el alcance de la presente invención.

Descripción de la forma preferida de ejecución

Con el fin de dar claridad a esta descripción, pero en modo alguno para limitarla, la descripción detallada se dividirá en las subsecciones siguientes:

I. Moléculas aisladas de ácido nucleico que codifican a 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 para detectar 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. Constructos de DNA que contienen una molécula de ácido nucleico de PCA3 y células que contienen estos constructos.

VII. Un anticuerpo que tiene afinidad de fijación con un polipéptido de PCA3 y un hibridoma que contiene el anticuerpo.

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

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

X. Exploración para el diagnóstico

XI. Tratamientos terapéuticos

XII. Animales no humanos transgénicos de PCA3

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

En una forma de ejecución, la presente invención se refiere a moléculas de ácido nucleico de PCA3 aisladas (purificadas). Con preferencia, la molécula de ácido nucleico de PCA3 contiene una secuencia de polinucleótido que es idéntica por lo menos en un 90% (con mayor preferencia, en un 95%, 96%, 97%, 98%, 99% o 100%) con una secuencia elegida entre el grupo formado por:

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

(b) una secuencia de nucleótidos que codifica a un polipéptido de PCA3 que contiene la secuencia completa de aminoácidos codificada por el clon de polinucleótido registrado en el depósito de la Centraal voor Schimmelcultures con el número de registro CBS 682.97;

5 (c) una secuencia de nucleótidos que codifica a un polipéptido de PCA3 que contiene la secuencia completa de aminoácidos codificada por el clon de polinucleótido registrado en el depósito de la Centraal voor Schimmelcultures con el número de registro CBS 100521; y

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

15 El pMB9 es un clon de cDNA de PCA3 que contiene los exones 1, 2, 3, 4a y 4b del gen PCA3. El pMB9 se ha depositado en la Phabagen Collection, University of Utrecht, Padualaan 8, 3584 CH Utrecht (que es una sección del Centraalbureau voor Schimmelcultures, Oosterstratt 1, código postal 273, 3740 AG Baarn) conforme a las cláusulas del Tratado Budapest en fecha 10 de abril de 1997, con el número de registro CBS 682.97.

20 El λ DD3.6 es un clon de cDNA de PCA3 cDNA que contiene los exones 3, 4a, 4b, 4c y 4d. El λ DD3.6 se ha depositado en la Phabagen Collection, University of Utrecht, Padualaan 8, 3584 CH Utrecht (que es una sección del Centraalbureau voor Schimmelcultures, Oosterstratt 1, código postal 273, 3740 AG Baarn) conforme a las cláusulas del Tratado Budapest en fecha 27 de marzo de 1998 con el número de registro CBS 100521.

25 En una forma preferida de ejecución, la molécula aislada de ácido nucleico contiene una secuencia de nucleótidos de PCA3 con una identidad o similitud mayor del 90% con la secuencia de nucleótidos definida en la SEQ ID NO: 1 (con preferencia mayor del 95%, 96%, 97%, 98%, 99% o 100%). En otra forma preferida de ejecución, la molécula aislada de ácido nucleico contiene la secuencia que codifica al PCA3 definida en la SEQ ID NO: 1. En otra forma de ejecución, la molécula aislada de ácido nucleico codifica a la secuencia de aminoácidos de PCA3 definida en la SEQ ID NO: 2 o SEQ ID NO: 7. En otra forma de ejecución más, la molécula aislada de ácido nucleico contiene una secuencia de nucleótidos de PCA3 que tiene una identidad o similitud mayor del 90% con la secuencia de nucleótidos definida en la SEQ ID NO: 6 (con preferencia mayor del 95%, 96%, 97%, 98%, 99% o 100%). En otra forma preferida
30 de ejecución, la molécula aislada de ácido nucleico contiene la secuencia que codifica al PCA3 definida en la SEQ ID NO: 6.

35 Se incluyen también dentro del alcance de esta invención los ácidos nucleicos aislados que contienen variantes de empalmes de cDNA de PCA3 o secuencias de polinucleótidos que tienen una identidad por lo menos del 90% con ellos, con preferencia una identidad por lo menos del 95% con ellos.

40 Teniendo en cuenta que virtualmente son posibles todas las combinaciones de exones, los ejemplos no limitantes de tales variantes de empalme incluyen a los ácidos nucleicos de PCA3 aislados, que contienen 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, ver 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, ver 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). Con preferencia, la molécula de ácido nucleico de PCA3 contiene una secuencia de polinucleótido que es idéntica por lo menos en un 90% (con mayor preferencia en un 95%, 96%, 97%, 98%, 99% o 100%) con una de las variantes de empalme recién descritas.

45 Se incluyen también dentro del alcance de esta invención los equivalentes funcionales de las moléculas aisladas de ácido nucleico aquí descritas y de sus derivados. Por ejemplo, las secuencias de ácido nucleico representadas en la SEQ ID NO: 1 o SEQ ID NO: 6 pueden alterarse con sustituciones, adiciones o deleciones, que proporcionan moléculas funcionalmente equivalentes.

50 Debido a la degeneración de las secuencias que codifican a nucleótidos, para la práctica de la presente invención podrán utilizarse otras secuencias de DNA que codifican sustancialmente a la misma secuencia de aminoácidos que se ha representado en la SEQ ID NO: 2 y SEQ ID NO: 7. Estas incluyen, pero no se limitan a: secuencias de nucleótidos que contienen la totalidad o porciones del ácido nucleico de PCA3 representado en las SEQ ID NO: 1, 3, 4 ó 6, que se alteran con la sustitución de diferentes codones, que codifican a un resto aminoácido funcionalmente equivalente dentro de la secuencia.

55 Además, la secuencia de ácido nucleico puede contener un nucleótido del extremo 5' y/o del extremo 3' de la fórmula de ácido nucleico representada en las SEQ ID NO: 1, 3, 4 ó 6 o de un derivado del mismo. A este respecto puede utilizarse cualquier nucleótido o polinucleótido, con la condición de que su adición, deleción o sustitución no altere sustancialmente la secuencia de aminoácidos de la SEQ ID NO: 2 o SEQ ID NO: 7, que se codifica con la secuencia de nucleótidos. Además, la molécula de ácido nucleico de la presente invención podrá tener, si fuera necesario, sitios de reconocimiento para la endonucleasa de restricción, que se añaden a su extremo 5' y/o 3'. Se incluyen, por tanto, en esta invención
60 todas las variaciones de la secuencia de nucleótidos del nucleótido de PCA3 que codifica a la secuencia y fragmentos de la misma permitidos por el código genético.

Es posible además suprimir codones o sustituir uno o más codones por codones que no sean codones degenerados con el fin de producir un polipéptido estructuralmente modificado, pero que tiene sustancialmente la misma utilidad o actividad que el polipéptido producido con una molécula no modificada de ácido nucleico. En la técnica se acepta que dos polipéptidos son funcionalmente equivalentes si lo son las dos moléculas de ácido nucleico que dan lugar a su producción, incluso cuando las diferencias entre las moléculas de ácido nucleico no guarden relación con la degeneración del código genético.

A. Aislamiento del ácido nucleico

En un aspecto de la presente invención se proporcionan moléculas aisladas de ácido nucleico que codifican a los polipéptidos que tienen secuencias de aminoácidos correspondientes al PCA3. En particular, la molécula de ácido nucleico puede aislarse a partir de una muestra biológica que contenga el RNA o del DNA del PCA3.

La molécula de ácido nucleico puede aislarse a partir de una muestra biológica que contenga el RNA del PCA3 aplicando técnicas de clonación de cDNA y una hibridación sustractiva. La molécula de ácido nucleico puede aislarse también a partir de una biblioteca de cDNA empleando una sonda homóloga.

La molécula de ácido nucleico puede aislarse a partir de una muestra biológica que contenga el DNA genómico o a partir de una biblioteca genómica. Las muestras biológicas idóneas incluyen, pero no se limitan a: organismos enteros, órganos, tejidos, sangre y células. El método para obtener la muestra biológica dependerá de la naturaleza de la muestra.

Los expertos ya saben que los genomas pueden presentar ligeras variaciones alélicas entre individuos. Por lo tanto, se acepta que la molécula aislada de ácido nucleico incluya también las variaciones alélicas, en el supuesto de que la secuencia sea un derivado funcional de la secuencia que codifica al PCA3. Si un alelo de PCA3 no codifica a una secuencia idéntica a la representada en la SEQ ID NO: 1 ó 6, podrá aislarse e identificarse como PCA3 aplicando las técnicas aquí descritas y en especial la técnica PCR para amplificar el gen apropiado con cebadores basados en las secuencias aquí descritas.

Los expertos ya saben que los organismos distintos de los humanos pueden contener también genes de PCA3 (por ejemplo, eucariotas; más en concreto: mamíferos, aves, peces y plantas; más específicamente: gorilas, monos rhesus y chimpancés). La invención incluye, pero no se limita a: moléculas de ácido nucleico de PCA3 aisladas de los organismos recién mencionados.

B. Síntesis de ácido nucleico

Las moléculas aisladas de ácido nucleico de la presente invención incluyen a las obtenidas por síntesis química. Por ejemplo, puede diseñarse una molécula de ácido nucleico que tenga la secuencia de nucleótidos que codifica al producto de expresión de un gen de PCA3 y, si fuera necesario, dividirse en fragmentos apropiados más pequeños. Se puede sintetizar, pues, un oligómero que corresponda a la molécula de ácido nucleico, o a cada uno de los fragmentos resultantes de la división. Estos oligonucleótidos sintéticos pueden prepararse, por ejemplo por el método del triéster de Matteucci y col., J. Am. Chem. Soc. 103, 3185-3191, 1981, o empleando un sintetizador automático de DNA.

Un oligonucleótido puede derivarse por síntesis o por clonación. Si fuera necesario, los extremos 5' de los oligómeros pueden fosforilarse empleando la quinasa de polinucleótido T4. Las hebras simples se someten a la acción de la quinasa antes de la fusión o para el marcado, esto puede lograrse empleando un exceso de enzima. Si la quinasa se emplea para marcar la muestra, entonces el ATP podrá contener radioisótopos de actividad muy específica. A continuación puede someterse el DNA oligómero a la fusión y ligación con la ligasa T4 u otra similar.

II. Polipéptidos de PCA3 purificados

Se describe también aquí un polipéptido purificado (con preferencia, sustancialmente puro) que tiene una secuencia de aminoácidos correspondiente al PCA3, o un derivado funcional del mismo. En una forma preferida de ejecución, el polipéptido tiene la secuencia de aminoácidos definida en la SEQ ID NO: 2 ó 7 o una mutante o una variación de especie de la misma, o que tiene una identidad por lo menos del 80% o que tiene una similitud por lo menos del 90% con la misma (con preferencia, una identidad por lo menos del 90%, 95%, 96%, 97%, 98%, o 99% o una similitud por lo menos del 95%, 96%, 97%, 98%, o 99% con la misma), o por lo menos 6 aminoácidos contiguos de la misma (con preferencia, por lo menos 10, 15, 20, 25 ó 50 aminoácidos contiguos de la misma).

También se describen aquí epítopes de PCA3. El epítipo de estos polipéptidos es un epítipo inmunogénico o antigénico. Un epítipo inmunogénico es aquella parte de la proteína que provoca la respuesta de un anticuerpo cuando la proteína entera es un inmunógeno. El epítipo antigénico es un fragmento de la proteína que puede desencadenar la respuesta de un anticuerpo. En la técnica ya se conocen los métodos para seleccionar fragmentos de epítopes antigénicos, véase Sutcliffe y col., Science 219, 660-666, 1983.

ES 2 324 503 T3

Los péptidos y polipéptidos que llevan epítopes antigénicos aquí descritos son útiles para provocar una respuesta inmune que reconoce específicamente a los polipéptidos. Los péptidos y polipéptidos que llevan epítopes antigénicos aquí descritos contienen por lo menos 7 aminoácidos (con preferencia, 9, 10, 12, 15 ó 20 aminoácidos) de las proteínas aquí descritas. Un ejemplo de péptido antigénico es el HTQEAQKEAQR (SEQ ID NO: 5).

Las variantes de secuencias de aminoácidos de PCA3 pueden obtenerse por mutaciones del DNA. Estas variantes incluyen, por ejemplo, deleciones de o inserciones en o sustituciones de restos dentro de la secuencia de aminoácidos definida en la SEQ ID NO: 2 ó 7. Puede realizarse también cualquier combinación de deleción, inserción y sustitución para llegar al constructo final, con la condición de que el constructo final posea la actividad deseada.

El sitio para la introducción de una variación de la secuencia de aminoácidos está predeterminado, mientras que la mutación de por sí no tiene que estar predeterminada. Por ejemplo, para optimizar la realización de una mutación en un sitio dado puede efectuarse una mutagénesis aleatoria en el codón o en la región diana y las variantes de PCA3 expresadas se exploran para encontrar la combinación óptima de la actividad deseada. Ya se conocen las técnicas para realizar mutaciones de sustitución en sitios predeterminados de una DNA que tenga una secuencia conocida, por ejemplo las mutagénesis específicas de sitio.

La preparación de una variante de PCA3 con arreglo a esto se lleva a cabo con preferencia mutagénesis de DNA específica de sitio, que codifica una variante preparada antes o una versión no variante de la proteína. La mutagénesis específica de sitio permite la producción de variantes de PCA3 mediante el uso de secuencias específicas de oligonucleótido que codifican la secuencia de DNA de la mutación deseada. En general, la técnica de la mutagénesis específica de sitio ya es conocida, y se ha descrito en publicaciones tales como Adelman y col., DNA 2, 183, 1983 y Ausubel y col. "Current Protocols in Molecular Biology", J. Wiley & Sons, Nueva York, NY, 1996.

Las deleciones en la secuencia de aminoácidos afectan en general de 1 a 30 restos, con mayor preferencia de 1 a 10 restos, y normalmente dichos restos son contiguos.

Las inserciones en la secuencia de aminoácidos incluyen las fusiones de amino y/o carboxilo terminales de un resto a polipéptidos de longitud esencialmente ilimitada, así como las inserciones dentro de la secuencia de restos que tienen uno o múltiples aminoácidos. Las inserciones dentro de una secuencia (es decir, inserciones dentro de la secuencia completa del PCA3) pueden añadir en general de 1 a 10 restos, con mayor preferencia de 1 a 5.

El tercer grupo de variantes es el formado por aquellas en las que por lo menos un resto aminoácido de la molécula del PCA3, y con preferencia solamente uno, se ha eliminado y en su lugar se ha insertado un resto diferente. Estas sustituciones se realizan con preferencia con arreglo a la siguiente tabla 1 cuando se desea modular finamente las características del PCA3.

(Tabla pasa a página siguiente)

ES 2 324 503 T3

TABLA 1

	resto original	sustituciones ilustrativas
5	Ala	gly; ser
	Arg	lys
10	Asn	gln, his
	Asp	glu
15	Cys	ser
	Gln	asn
20	Glu	asp
	Gly	ala; pro
25	His	asn; gln
	Ile	leu; val
	Leu	ile; val
30	Lys	arg; gln; glu
	Met	leu; tyr; ile
35	Phe	met; leu; tyr
	Ser	thr
40	Thr	ser
	Trp	tyr
45	Tyr	trp; phe
	Val	ile; leu

Se hacen cambios sustanciales en la identidad funcional o inmunológica seleccionando sustituciones que sean menos conservadoras que las de la tabla 1, es decir, seleccionando restos que difieran de modo más significativa en su efecto de mantener (a) la estructura del esqueleto del polipéptido en la zona de la sustitución, por ejemplo, adoptando la conformación de hoja o de hélice, (b) la carga o el carácter hidrófobo de la molécula en el sitio deseado, o (c) el tamaño o volumen de la cadena lateral. Las sustituciones que se esperan en general son aquellas, en las que (a) la glicina y/o la prolina se sustituye por otro aminoácido o se suprime o se inserta; (b) se sustituye un resto hidrófilo, p.ej., serilo o treonilo, por un resto hidrófilo, p.ej., leucilo, isoleucilo, fenilalanilo, valilo o alanilo; (c) se sustituye un resto cisteína por cualquier otro resto; (d) se sustituye un resto que tiene una cadena lateral electropositiva, p.ej., lisilo, arginilo o histidilo, por un resto que tenga carga electronegativa, p.ej., glutamilo o aspartilo; o (e) se sustituye un resto que tiene la cadena lateral voluminosa, p.ej., fenilalanina, por otro que no tenga una cadena lateral de este tipo, p.ej. la glicina.

Algunas deleciones, inserciones y sustituciones no se espera que produzcan cambios radicales en las características del PCA3. Sin embargo, cuando es difícil predecir el alcance exacto de una sustitución, deleción o inserción por anticipado, los expertos ya saben que se puede evaluar dicho alcance mediante ensayos exploratorios rutinarios. Por ejemplo, puede conseguirse una variante típica por mutagénesis específica de sitio del ácido nucleico que codifica al PCA3 nativo, la expresión de la variante de ácido nucleico en cultivo celular recombinante y opcionalmente la purificación del cultivo celular, por ejemplo, por adsorción de inmunoafinidad en una columna (para absorber la variante fijándola por lo menos sobre un epítipo inmune restante). La actividad del lisado celular o de la variante de molécula de PCA3 purificada se explora en un ensayo adecuado que permita determinar la característica deseada. Por ejemplo,

se mide un cambio del carácter inmunológico de la molécula de PCA3, por ejemplo la afinidad con un anticuerpo determinado, mediante un inmunoensayo de tipo competitivo. Los cambios en la actividad de inmunomodulación se miden con ensayos apropiados. Las modificaciones de estas propiedades de las proteínas, por ejemplo la estabilidad redox o térmica, el carácter hidrófobo, la susceptibilidad de sufrir una degradación proteolítica o la tendencia a agregarse con vehículos o en multímeros pueden ensayar por métodos que los expertos conocen bien.

Se puede emplear un gran número de metodología ya conocidas de la técnica para obtener el péptido aquí descrito. En una forma de ejecución se purifica el péptido a partir de tejidos o células que producen el péptido de manera natural. Como alternativa, los fragmentos de ácido nucleico aislados recién descritos pueden emplearse para expresar la proteína de PCA3 en cualquier organismo. Las muestras empleadas en la presente invención incluyen a células, extractos proteicos o extractos de membranas de células, o líquidos biológicos. La muestra puede variar según el formato del ensayo, el método de detección y la naturaleza de los tejidos, células o extractos recogidos en la muestra.

Puede utilizarse cualquier organismo como fuente del péptido aquí descrito, en el supuesto de que el organismo fuente contenga de modo natural dicho péptido. Tal como se emplea aquí, "organismo fuente" indica el organismo original a partir del cual se ha derivado la secuencia de aminoácidos de la subunidad, con independencia del organismo en el que se expresa la subunidad o a fin de cuenta a partir del que se ha aislado.

Los expertos pueden adoptar fácilmente métodos conocidos para aislar proteínas con el fin de obtener el péptido libre de los contaminantes naturales. Estos métodos incluyen, pero no se limitan a: inmunocromatografía, cromatografía de exclusión de tamaños, HPLC, cromatografía de intercambio iónico y cromatografía de inmuno-afinidad.

En una forma preferida de ejecución, los procedimientos de purificación consisten en la cromatografía de intercambio iónico y cromatografía de exclusión de tamaños. Puede emplearse cualquiera de las muchas resinas de intercambio iónico conocidas, por ejemplo la monoQ, Sepharose Q, Macro-prepQ, AG1-X2, o HQ. Los ejemplos de de resinas idóneas para la exclusión de tamaños incluyen, pero no se limitan a: Superdex 200, Superose 12, y Sephacryl 200. La elución puede realizarse con soluciones acuosas de cloruro potásico o de cloruro sódico, de concentraciones comprendidas entre 0,01 M y 2,0 M.

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

En otra forma de ejecución, la presente invención se refiere a una sonda de ácido nucleico para la detección específica de la presencia de ácido nucleico de PCA3 en una muestra que contiene las moléculas de ácido nucleico antes descritas o por lo menos un fragmento de las mismas que en condiciones restrictivas se fije sobre el ácido nucleico del PCA3.

En una forma preferida de ejecución, la presente invención se refiere a una sonda de ácido nucleico aislada que contiene de 10 a 1000 nucleótidos (con preferencia, 10-500, 10-100, 10-50, 10-35, 20-1000, 20-500, 20-100, 20-50 o 20-35), que se hibrida con preferencia con el RNA o DNA que codifica al PCA3 o con un gen de PCA3, pero no con un RNA o DNA que no sea afín al PCA3, dicha sonda de ácido nucleico es o es complementaria de una secuencia de nucleótidos formada por lo menos por 10 nucleótidos consecutivos (con preferencia 15, 18, 20, 25 ó 30) de la molécula de ácido nucleico que contiene una secuencia de polinucleótido que es idéntica por lo menos en un 90% con una secuencia elegida entre el grupo formado por:

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

(b) una secuencia de nucleótidos que codifica al polipéptido de PCA3 que contiene la secuencia completa de aminoácidos codificada por el clon polinucleótido entregado en depósito al Centraal voor Schimmelcultures con el número de registro CBS 682.97, respectivamente;

(c) una secuencia de nucleótidos codifica al polipéptido de PCA3 que contiene la secuencia completa de aminoácidos codificada por el clon polinucleótido entregado en depósito al Centraal voor Schimmelcultures con el número de registro CBS 100521, respectivamente;

(d) una secuencia de nucleótidos que codifica al gen PCA3 que contiene la secuencia de nucleótidos definida en la SEQ ID NO: 1, 3, 4 ó 6;

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

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

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

(h) una secuencia de nucleótidos que se ha descrito en páginas anteriores.

ES 2 324 503 T3

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

- 5 Las secuencias complementarias son también conocidas como ácidos nucleicos antisentido cuando contienen secuencias que son complementarias de la hebra codificadora.

Los ejemplos de sondas específicas de ácido nucleico que pueden utilizarse en la presente invención se recogen en la siguiente tabla 2.

TABLA 2

Sondas de ácido nucleico

	tamaño (nº de bases)	nucleótidos
exón 1	98	1-98 de la SEQ ID NO: 1
exón 2	165	99-263 de la SEQ ID NO: 1
exón 3	183	264-446 de la SEQ ID NO: 1
exón 4a	539	447-985 de la SEQ ID NO: 1
exón 4b	1052	986-2037 de la SEQ ID NO: 1
sonda 1	20	1-20 de la SEQ ID NO: 1
sonda 2	30	1-30 de la SEQ ID NO: 1
sonda 3	40	1-40 de la SEQ ID NO: 1
sonda 4	20	381-400 de la SEQ ID NO: 1
sonda 5	30	381-410 de la SEQ ID NO: 1
sonda 6	20	401-420 de la SEQ ID NO: 1
sonda 7	30	401-430 de la SEQ ID NO: 1
sonda 8	20	511-530 de la SEQ ID NO: 1
sonda 9	30	501-530 de la SEQ ID NO: 1
sonda 10	20	77-98 de la SEQ ID NO: 1
sonda 11	20	99-118 de la SEQ ID NO: 1
sonda 12	20	244-263 de la SEQ ID NO: 1
sonda 13	20	264-283 de la SEQ ID NO: 1
sonda 14	20	427-446 de la SEQ ID NO: 1

	tamaño (n° de bases)	nucleótidos
sonda 15	20	447-466 de la SEQ ID NO: 1
exón 1	120	1-120 de la SEQ ID NO: 6
exón 2	165	121-285 de la SEQ ID NO: 6
exón 3	183	286-468 de la SEQ ID NO: 6
exón 4a	539	469-1007 de la SEQ ID NO: 6
exón 4b	1059	1008-2066 de la SEQ ID NO: 6
exón 4c	556	2067-2622 de la SEQ ID NO: 6
exón 4d	960	2623-3582 de la SEQ ID NO: 6

Obviamente, los expertos ya saben que se puede diseñar un gran número de sondas adicionales a partir de la misma región o de otra región de la SEQ ID NO: 1 y también de la SEQ ID NO: 6 o de otras secuencias de la presente invención.

La sonda de ácido nucleico puede utilizarse para explorar una biblioteca adecuada de DNA cromosómicos o de cDNA con los métodos usuales de hibridación para obtener otra molécula de ácido nucleico de la presente invención. Una biblioteca adecuada de DNA cromosómicos o de cDNA puede prepararse a partir de células apropiadas aplicando métodos ya reconocidos de la técnica (véase Molecular Cloning, A Laboratory Manual, segunda edición, coordinado por Sambrook, Fritsch & Maniatis, Cold Spring Harbor Laboratory, 1989).

Como alternativa, la síntesis química se efectúa con el fin de obtener sondas de ácido nucleico que tengan secuencias de nucleótidos que correspondan a las porciones N-terminal y C-terminal de la secuencia de aminoácidos del PCA3. Por lo tanto, las sondas de ácido nucleico sintetizadas pueden utilizarse como cebadores en una reacción en cadena de la polimerasa (PCR) realizada con arreglo a técnicas PCR reconocidas, esencialmente según los PCR Protocols, A Guide to Methods and Applications, coordinado por Michael y col., Academic Press, 1990, empleando una biblioteca apropiada de cromosomas, de cDNA o de líneas celulares, para obtener el fragmento de la presente invención.

Los expertos pueden diseñar fácilmente estas sondas basadas en la secuencia aquí descrita, aplicando métodos de alineamiento y análisis de secuencia computerizados ya conocidos en la técnica (véase Molecular Cloning, A Laboratory Manual, segunda edición, coordinado por Sambrook, Fritsch & Maniatis, Cold Spring Harbor Laboratory, 1989).

Las sondas de hibridación de la presente invención pueden marcarse por técnicas estándar, por ejemplo con radiomarcadores, con marcadores enzimáticos, con marcadores fluorescentes, con marcadores de biotina-avidina, con marcadores quimioluminiscentes y similares. Después de la hibridación, las sondas pueden visualizarse empleando métodos ya conocidos.

Las sondas de ácido nucleico de la presente invención incluyen a las sondas de RNA y también a las de DNA, dichas se generan por métodos ya conocidos de la técnica.

En una forma de ejecución del método recién descrito se inmoviliza una sonda de ácido nucleico sobre un soporte sólido.

Los ejemplos de tales soportes sólidos incluyen, pero no se limitan a: plásticos, por ejemplo el policarbonato, hidratos de carbono complejos, por ejemplo la agarosa y la sefarosa, las resinas acrílicas, por ejemplo la poliácridamida y las bolas de látex. En la técnica ya se conocen los métodos para fijar las sondas de ácido nucleico sobre estos tipos de soportes sólidos.

Las muestras a ensayar, idóneas para los métodos de detección de ácido nucleico de la presente invención, incluyen, por ejemplo, células o extractos de ácido nucleico de células o líquidos biológicos. La muestra empleada en los métodos recién descritos puede variar según el formato del ensayo, el método de detección y la naturaleza de los tejidos, células o extractos a ensayar. Los expertos ya conocen los métodos de preparación de extractos de ácido

nucleico de células y pueden adaptarse fácilmente para la obtención de una muestra que sea compatible con el método utilizado.

IV. *Método para detectar la presencia de ácido nucleico de PCA3 en una muestra*

En otra forma de ejecución, la presente invención se refiere a un método para detectar la presencia de ácido nucleico de PCA3 en una muestra, que consiste en: a) poner en contacto la muestra con la sonda de ácido nucleico antes descrita, en condiciones específicas de hibridación, de manera que se produzca dicha hibridación, y b) detectar la presencia de la sonda fijada sobre la molécula de ácido nucleico de PCA3. Los expertos ya saben seleccionar la sonda de ácido nucleico según los métodos ya conocidos de la técnica que se han descrito en páginas anteriores. Las muestras a analizar incluyen, pero no se limitan a: muestras de RNA o DNA de tejido humano.

V. *Kit para detectar la presencia de ácido nucleico de PCA3 en una muestra*

Se describe también un kit para detectar la presencia de ácido nucleico de PCA3 en una muestra que consta por lo menos de un recipiente que contiene la sonda de ácido nucleico recién descrita. En una forma preferida de ejecución, el kit consta además de otros recipientes que contienen uno o más de los siguientes: reactivos de lavado y reactivos capaces de detectar la presencia de sonda de ácido nucleico fijada. Los ejemplos de reactivos de detección incluyen, pero no se limitan a: sondas radiomarcadas, sondas marcadas con enzimas (peroxidasa de rábano rusticano, fosfatasa alcalina), y sondas marcadas por afinidad (biotina, avidina o estreptavidina).

En detalle, un kit dividido en compartimentos incluye cualquier kit, en el que los reactivos están alojados en recipientes separados. Estos recipientes (contenedores) incluyen recipientes pequeños de vidrio, recipientes de plástico o tiras de plástico o de papel. Estos recipientes impiden la transferencia de los reactivos desde un compartimento a otro compartimento, de modo que las muestras y reactivos no se contaminen entre sí y los agentes o soluciones de cada recipiente puedan añadirse de modo cuantitativo de un compartimento a otro. Estos recipientes incluirán un recipiente, en el que se depositará la muestra a ensayar, un recipiente para contener la sonda o los cebadores empleados en el ensayo, recipientes para contener los reactivos de lavado (por ejemplo solución salina tamponada con fosfato, tampones Tris y similares), y recipientes para contener los reactivos empleados para detectar la sonda hibridada, el anticuerpo fijado, el producto amplificado o similares.

Los expertos reconocerán fácilmente que las sondas de ácido nucleico descritas en la presente invención pueden incorporarse fácilmente a uno de los formatos convencionales de kits, que dichos expertos conocen perfectamente.

VI. *Constructos de DNA que contienen una molécula de ácido nucleico de PCA3 y células que contienen estos constructos*

En otra forma de ejecución, la presente invención se refiere a una molécula recombinante de DNA que consta de: 5'-3', un promotor eficaz para iniciar la transcripción en una célula hospedante y moléculas de ácido nucleico antes descritas. En otra forma de ejecución, la presente invención se refiere a una molécula recombinante de DNA que consta de un vector y una molécula de ácido nucleico antes descrita.

En otra forma de ejecución, la presente invención se refiere a una molécula de ácido nucleico que contiene una región de control de transcripción funcional de la célula, una secuencia complementaria de una secuencia de RNA que codifica a una secuencia de aminoácidos correspondiente al polipéptido antes descrito y una región de terminación de la transcripción funcional de la célula.

Con preferencia, las moléculas recién descritas son moléculas de DNA purificadas y/o son aisladas.

En otra forma de ejecución, la presente invención se refiere a una célula u organismo no humano, que contiene una molécula de ácido nucleico antes descrita.

En otra forma de ejecución, el péptido se purifica a partir de células que se han alterado para que expresan al péptido.

Tal como se emplea aquí, se dice que una célula se ha "alterado para expresar un péptido deseado" cuando dicha célula, mediante manipulación genética, se ha generado para producir una proteína que normalmente no produce o que la célula normalmente produce en cantidades bajas. Los expertos sabrán adaptar fácilmente los procedimientos para introducir y expresar secuencias genómicas, de cDNA o sintéticas ya sea en células eucariotas, ya sea en procariotas.

Una molécula de ácido nucleico, por ejemplo el DNA, se dice que es "capaz de expresar" un polipéptido si contiene las secuencias de nucleótidos que contienen información reguladora de la transcripción y de la traducción y tales secuencias y tales secuencias están "unidas operativamente" a secuencias de nucleótidos que codifican al polipéptido. Una unión operativa es una unión, en la que las secuencias reguladoras de DNA y la secuencia de DNA destinada a expresarse están conectadas de tal manera que permitan la expresión de la secuencia genética. La naturaleza precisa de las regiones reguladoras requeridas para la expresión de la secuencia genética pueden variar de un organismo a otro, pero en general deberán incluir una región de promotor que, en los procariotas, contiene tanto al promotor (que dirige

el inicio de la transcripción de RNA) como a las secuencias de DNA que, una vez transcritas al RNA, señalarán el inicio de la síntesis. Estas regiones incluirán normalmente aquellas secuencias no codificadoras de 5' que intervienen en el inicio de la transcripción y traducción, por ejemplo secuencia TATA (TATA box), secuencia de casquete, la secuencia CAAT y similares.

Si se desea, la región no codificadora 3' de la secuencia que codifica al PCA3 puede obtenerse por los métodos recién descritos. Esta región puede guardarse en lo que respecta a sus secuencias reguladoras de la terminación de la transcripción, por ejemplo la terminación y la poliadenilación. Por lo tanto, guardando la región 3' naturalmente contigua a la secuencia de DNA que codifica al gen PCA3, se pueden generar señales de terminación de la transcripción. Si las señales de terminación de la transcripción no son funcionalmente satisfactorias para la expresión en la célula hospedante, entonces podrá sustituirse la región 3' funcional de la célula hospedante.

Se dice que dos secuencias de DNA (por ejemplo una secuencia de región de promotor y una secuencia codificadora de PCA3) están unidas operativamente si la naturaleza de la unión entre las dos secuencias de DNA (1) no produce la introducción de una mutación con desplazamiento de marco, (2) no interfiere en la capacidad de la secuencia de la región de promotor para dirigir la transcripción de una secuencia codificadora de PCA3, o (3) no interfiere en la capacidad de la secuencia que codifica al PCA3 de transcribir por acción de la secuencia de la región del promotor. Por lo tanto, una región de promotor estaría unida operativamente con una secuencia de DNA si el promotor fuera capaz de efectuar la transcripción de dicha secuencia de DNA.

La presente invención abarca la expresión de la secuencia que codifica al PCA3 (o un derivado funcional de la misma) en células procariotas o eucariotas. Los hospedantes procariotas son, en general, los más eficaces y convenientes para la producción de proteínas recombinantes y, por ello, son preferidos para la expresión de la secuencia que codifica al PCA3.

Los procariotas se representan muy a menudo por varias cepas de *E. coli*. Sin embargo, pueden utilizarse también otras cepas microbianas, incluidas otras cepas bacterianas. En los sistemas procariotas pueden utilizarse vectores plásmidos que contienen sitios de replicación y secuencias de control derivadas de una especie compatible con el hospedante. Los ejemplos de vectores plásmidos apropiados incluyen al pBR322, pUC18, pUC19, pUC118, pUC119 y similares; los vectores fagos o bacteriófagos idóneos incluyen al λ gt10, λ gt11 y similares; y los vectores víricos idóneos incluyen al pMAM-neo, pKRC y similares. Con preferencia, el vector elegido de la presente invención tiene la capacidad de replicarse en la célula hospedante elegida.

Los hospedantes procariotas reconocidos incluyen a las bacterias del tipo *E. coli*, *Bacillus*, *Streptomyces*, *Pseudomonas*, *Salmonella*, *Serratia* y similares. Sin embargo, en estas condiciones, el péptido no se glucosilará. El hospedante procariota tiene que ser compatible con el replicón y las secuencias de control del plásmido de expresión.

Para expresar el PCA3 en una célula procariota es necesario unir operativamente la secuencia que codifica al PCA3 con el promotor procariota funcional. Tales promotores pueden ser constitutivo o, con mayor preferencia, regulables (es decir, inducibles o desrepresibles). Los ejemplos de promotores constitutivos incluyen al promotor int del bacteriófago λ , el promotor bla de la secuencia del gen de la β -lactamasa del pBR322 y el promotor CAT de la secuencia del gen del cloranfenicol-acetil-transferasa del pBR325 y similares. Los ejemplos de promotores procariotas inducibles incluyen los principales promotores de derecha e izquierda del bacteriófago λ (P_L y P_R), los promotores trp, recA, lacZ, lacI y gal de la *E. coli*, la α -amilasa (Ulmanen y col., J. Bacteriol. 162, 176-182, 1985) y los promotores específicos de ζ -28 del *B. subtilis* (Gilman y col., 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 los promotores de *Streptomyces* (Ward y col., Mol. Gen. Genet. 203, 468-478, 1986). Los promotores procariotas han sido objeto de revisión en 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 correcta en una célula procariota puede requerir también la presencia de un sitio de unión a ribosoma en posición anterior (upstream) dentro de la secuencia que codifica a la secuencia del gen. Estos sitios de unión a ribosoma se han descrito, por ejemplo, en Gold y col. (Ann. Rev. Microbiol. 35, 365-404, 1981).

La selección de las secuencias de control, vectores de expresión, métodos de transformación y similares, dependerá del tipo de célula hospedante empleada para expresar el gen. Tal como se emplea aquí, "célula", "línea celular" y "cultivo celular" pueden usarse indistintamente y todas estas designaciones incluyen a la progenie. Por lo tanto, los términos "transformantes" o "células transformadas" incluyen a la célula primaria en cuestión y a los cultivos derivados de ella, prescindiendo del número de transferencias. Se da también por supuestos que toda la progenie no tiene que ser necesariamente idéntica en contenido de DNA, debido a las mutaciones deliberadas o fortuitas.

Sin embargo, tal como se ha definido, la progenie mutante tiene la misma funcionalidad que la célula transformada original. Las células hospedantes que pueden emplearse para los sistemas de expresión de la presente invención no están limitadas de forma estricta, excepto que deberán ser idóneas para el uso en la expresión del péptido PCA3 de interés. Los hospedantes idóneos incluyen a las células eucariotas.

Los hospedantes eucariotas preferidos incluyen, por ejemplo, a la levadura, los hongos, las células de insectos, las células de mamíferos ya sea "in vivo", ya sea en cultivo de Tejido. Las células de mamífero preferidas incluyen a las

células HeLa, a las células de origen fibroblástico, por ejemplo las VERO o CHO-K1, o las células de origen linfóide y sus derivados.

Además, las células vegetales pueden emplearse también como hospedantes; se dispone de secuencias de control compatibles con células vegetales, por ejemplo el virus del mosaico de la coliflor 35S y 19S y las secuencias de promotor de nopalina-sintasa y de las señales de poliadenilación.

Otro hospedante preferido es una célula de insecto, por ejemplo larvas de *Drosophila*. Empleando las células de insecto como hospedantes puede utilizarse el promotor de deshidrogenasa de alcohol de la *Drosophila*, véase Rubin, Science 240, 1453-1459, 1988. Como alternativa, pueden diseñarse por ingeniería genética los vectores de baculovirus para que expresen grandes cantidades de PCA3 en células de insectos (Jasny, Science 238, 1653, 1987; Miller y col., en: Genetic Engineering (1986), coordinado por Setlow, J.K. y col., editorial Plenum, vol. 8, pp. 277-297).

Las diferentes células hospedantes tienen mecanismos característicos y específicos para el proceso de traducción y post-traducción y para la modificación (p.ej. glucosilación, descomposición) de las proteínas. Las líneas celulares o sistemas de hospedantes apropiados podrán elegirse para asegurar la modificación y procesamiento deseados de la proteína foránea expresada.

Puede utilizarse cualquiera de una serie de sistemas de expresión de secuencias de genes de levadura que incorporan elementos promotor y de terminación a partir de secuencias genéticas expresadas activamente, que codifican a las enzimas glucolíticas. Estas enzimas se producen en grandes cantidades cuando se cultivan las levaduras en medios ricos en glucosa. Las secuencias genéticas glucolíticas conocidas pueden proporcionar también señales muy eficaces de control de la transcripción.

La levadura proporciona ventajas sustanciales por el hecho de que puede efectuar también modificaciones peptídicas post-traduccionales. Existe un gran número de estrategias de DNA recombinante que utilizan secuencias de promotor fuerte y un número de elevado de copias de plásmidos, que pueden utilizarse para la producción de las proteínas deseadas en la levadura. La levadura reconoce secuencia líder en productos de secuencia genética de mamífero clonada y segrega péptidos que llevan secuencias líder (es decir, pre-péptidos). Para un hospedante mamífero, se dispone de diversos sistemas de vectores posibles para la expresión del PCA3.

Puede emplearse una gran variedad de secuencias reguladoras de la transcripción y de la traducción, en función de la naturaleza del hospedante. Las señales reguladoras de la transcripción y de la traducción pueden derivarse de fuentes víricas, por ejemplo adenovirus, del virus del papiloma bovino, de virus de simios o similares, en los que las señales reguladoras están asociadas con una secuencia genética particular, que tiene un nivel de expresión alto. Como alternativa, pueden emplearse promotores procedentes de productos de expresión de mamíferos, por ejemplo la actina, el colágeno, la miosina y similares. Las señales reguladoras del inicio de la transcripción pueden elegirse de modo que permitan la represión o la activación, con lo cual podrá modularse la expresión de las secuencias genéticas. Son también de interés las señales reguladoras que son sensibles a la temperatura, de modo que variando la temperatura puede reprimirse o iniciarse la expresión o son susceptibles de regulación química (por ejemplo un metabolito).

Tal como se ha debatido antes, la expresión del PCA3 en hospedantes eucariotas requiere el uso de regiones reguladoras eucariotas. Estas regiones incluyen por lo general una región de promotor suficiente para dirigir el inicio de la síntesis del RNA. Los promotores eucariotas preferidos incluyen, por ejemplo, al promotor de la secuencia del gen de la metalotioneína I del ratón (Hamer y col., J. Mol. Appl. Gen. 1, 273-288, 1982); al promotor TK del herpes virus (McKnight, Cell 31, 355-365, 1982); al promotor temprano SV40 (Benoist y col., Nature (Londres) 290, 304-310, 1981); al promotor de la secuencia genética gal4 de la levadura (Johnston y col., Proc. Natl. Acad. Sci. (USA) 79, 6971-6975, 1982; Silver y col., Proc. Natl. Acad. Sci. (USA) 81, 5951-5955, 1984) y al promotor del gen inmediato-temprano del CMV (Thomsen y col., Proc. Natl. Acad. Sci. (USA) 81, 659-663, 1984).

Como ya es sabido en general, la traducción del mRNA eucariota se inicia en el codón que codifica a la primera metionina. Por esta razón es preferible asegurar que el engarce entre el promotor eucariota y la secuencia que codifica al PCA3 no contenga ningún codón que pudiera intervenir y sea capaz de codificar a la metionina (es decir, AUG). La presencia de tales codones tiene como resultado la formación de una proteína de fusión (si el codón AUG se halla en el mismo marco de lectura que la secuencia que codifica al PCA3) o una mutación de desplazamiento de marco (si el codón AUG no se halla en el mismo marco de lectura que la secuencia que codifica al PCA3).

Una molécula de ácido nucleico de PCA3 y un promotor unido operativamente pueden introducirse en una célula receptora procariota o eucariota ya sea como molécula de DNA (o de RNA) no replicante, que puede ser una molécula lineal, ya sea, con mayor preferencia, una molécula circular covalente cerrada. Dado que tales moléculas son incapaces de replicación autónoma, la expresión del gen puede tener lugar mediante la expresión transitoria de la secuencia introducida. Como alternativa, la expresión permanente puede ocurrir mediante la integración de la secuencia de DNA introducida en el cromosoma hospedante.

En una forma de ejecución, se emplea un vector capaz de integrar las secuencias genéticas deseadas en el cromosoma de la célula hospedante. Las células que han integrado de modo estable el DNA introducido en sus cromosomas pueden seleccionarse introduciendo también uno o más marcadores que permitan la selección de células hospedantes que contengan el vector de expresión. Los marcadores pueden proporcionar prototrofia con el hospedante auxotrófico,

resistencia a biocidas, p.ej. a antibióticos o a metales pesados, por ejemplo el cobre, y similares. La secuencia del gen marcador seleccionable puede unirse directamente a las secuencias genéticas de DNA a expresar o introducirse en la misma célula por co-transfección. Es posible que se requieran elementos adicionales para la síntesis óptima del mRNA de proteína de unión de cadena simple. Estos elementos pueden incluir señales de empalme así como promotores de transcripción, secuencias de señal de intensificador y señales de terminación. Los vectores de expresión de cDNA que incorporan estos elementos incluyen a los descritos por Okayama, Molec. Cell. Biol. 3, 280, 1983.

En una forma preferida de ejecución, la molécula de ácido nucleico introducida se incorporará a un vector plásmido o vírico capaz de replicación autónoma en el hospedante receptor. Para este fin puede emplearse uno cualquiera dentro de una gran variedad de vectores. Los factores importantes para elegir un vector plásmido o vírico concreto son: la facilidad con que se reconocen y seleccionan las células receptoras que contienen el vector entre las células receptoras que no contienen al vector; el número de copias del vector que se desean en un hospedante concreto; y si es deseable que sea capaz de transportar (shuttle) el vector entre células hospedantes de diferentes especies. Los vectores procariotas preferidos incluyen a los plásmidos, por ejemplo los que son capaces de replicación en *E. coli* (por ejemplo el pBR322, ColE1, pSC101, pACYC 184, π VX). Estos plásmidos se han descrito por ejemplo en Sambrook (véase Molecular Cloning: A Laboratory Manual, segunda edición, coordinado por Sambrook, Fritsch & Maniatis, Cold Spring Harbor Laboratory, 1989). Los plásmidos de bacilos incluyen al pC194, pC221, pT127 y similares. Estos plásmidos se han descrito en Gryczan (en: The Molecular Biology of the Bacilli, Academic Press, NY, 1982, pp. 307-329). Los plásmidos adecuados de *Streptomyces* incluyen al pIJ101 (Kendall y col., J. Bacteriol. 169, 4177-4183, 1987) y a los bacteriófagos de *Streptomyces* por ejemplo el ϕ C31 (Chater y col., en: Sexto Simposio Internacional de Biología de Actinomicetales, Akademiai Kiado, Budapest, Hungría, pp. 45-54, 1986). Los plásmidos de *Pseudomonas* se han revisado en John y col. (Rev. Infect. Dis. 8, 693-704, 1986) y en Izaki (Jpn. J. Bacteriol. 33, 729-742, 1978).

Los plásmidos eucariotas preferidos incluyen, por ejemplo, al BPV, vacunas, SV40, círculos de 2 micras y similares, o sus derivados. Estos plásmidos son bien conocidos en la técnica (Botstein y col., 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 y col., J. Clin. Hematol. Oncol. 10, 39-48, 1980; Maniatis, en: Cell Biology: A Comprehensive Treatise, vol. 3, Gene Sequence Expression, Academic Press, NY, pp. 563-608, 1980).

Una vez se ha preparado para la expresión el vector o la molécula de ácido nucleico que contiene el o los constructos, el o los constructos de DNA pueden introducirse en una molécula hospedante apropiada por cualquiera de los muchos métodos apropiados, es decir, por transformación, transfección, conjugación, fusión de protoplastos, electroporación, tecnología de cañones de partículas, precipitación con fosfato cálcico, microinyección directa y similares. Después de la introducción del vector, se cultivan las células receptoras en un medio selectivo, que selecciona las células que contienen al vector para que crezcan. La expresión de la o de las moléculas de gen clonadas se traduce en la producción del PCA3. Esta tiene lugar en las células transformadas como tales o después de la inducción de estas células a diferenciarse (por ejemplo, por administración del bromodesoxiuracilo a las células de neuroblastoma o similares).

VII. Un anticuerpo que tiene afinidad de fijación con un polipéptido de PCA3 y un hibridoma que contiene el anticuerpo

Se describe también aquí un anticuerpo que tiene capacidad de unión específica con un polipéptido de PCA3 ya descrito antes, o con un polipéptido de PCA3 que se une a un fragmento del mismo. Un anticuerpo se une específicamente a un polipéptido de PCA3 o a un fragmento de fijación del mismo si no se une a polipéptidos que no sean de PCA3. Los que se unen selectivamente al PCA3 se eligen para métodos que pueden incluir, pero no se limitan a ellos, el análisis de la expresión alterada del PCA3 en un tejido que contenga al PCA3.

Las proteínas de PCA3 aquí descritas pueden utilizarse en una gran variedad de procedimientos y métodos, por ejemplo para la generación de anticuerpos, para la identificación de composiciones farmacéuticas y para el estudio de la interacción DNA/proteína.

El PCA3 péptido aquí descrito puede utilizarse para producir anticuerpos o hibridomas. Los expertos reconocerán fácilmente que si se desea un anticuerpo, tal péptido podría generarse del modo aquí descrito y emplearse como inmunógeno.

Los anticuerpos aquí descritos incluyen a los anticuerpos monoclonales y policlonales, así como a fragmentos de estos anticuerpos. Se describen además anticuerpos de cadena simple. Los fragmentos de anticuerpo que contienen el idiotipo de la molécula pueden generarse por técnicas conocidas. Por ejemplo, estos fragmentos incluyen, pero no se limitan a: el fragmento F(ab')₂; los fragmentos Fab', los fragmentos Fab y los fragmentos Fv.

Son de un interés especial los anticuerpos contra el PCA3 que se producen en los seres humanos o se "humanizan" (es decir, no son inmunogénicos para el ser humano) por tecnología recombinante o de otro tipo. Los anticuerpos humanizados pueden producirse por ejemplo reemplazando una porción inmunogénica de un anticuerpo por la correspondiente porción no inmunogénica (es decir, anticuerpos quiméricos) (Robinson, R.R. y col., publicación de patente internacional PCT/US 86/02269; Akira, K. y col., solicitud de patente europea 184,187; Taniguchi, M., solicitud de

patente europea 171,496; Morrison, S.L. y col., solicitud de patente europea 173,494; Neuberger, M.S. y col., solicitud PCT WO 86/01533; Cabilly, S. y col., solicitud de patente europea 125,023; Better, M. y col., Science 240, 1041-1043, 1988; Liu, A.Y. y col., Proc. Natl. Acad. Sci. USA 84, 3439-3443, 1987; Liu, A.Y. y col., J. Immunol. 139, 3521-3526, 1987; Sun, L.K. y col., Proc. Natl. Acad. Sci. USA 84, 214-218, 1987; Nishimura, Y. y col., Canc. Res. 47, 999-1005, 1987; Wood, C.R. y col., Nature 314, 446-449, 1985; Shaw y col., J. Natl. Cancer Inst. 80, 1553-1559, 1988). Una revisión general de los anticuerpos quiméricos “humanizados” se encontrará en Morrison, S.L. (Science 229, 1202-1207, 1985) y en Oi, V.T. y col., BioTechniques 4, 214, 1986). Los anticuerpos “humanizados” idóneos pueden producirse también, como alternativa, por CDR o sustitución CEA (Jones, P.T. y col., Nature 321, 552-525, 1986; Verhoeyan y col., Science 239, 1534, 1988; Beidler, C.B. y col., J. Immunol. 141, 4053-4060, 1988).

Se describe también aquí un hibridoma que produce el anticuerpo monoclonal antes descrito. Un hibridoma es una línea celular inmortalizada que es capaz de secretar un anticuerpo monoclonal específico.

En general, los métodos para obtener anticuerpos monoclonales e hibridomas son bien conocidas en la técnica (Campbell, “Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology”, Elsevier Science Publishers, Amsterdam, Holanda (1984); St. Groth y col., J. Immunol. Methods 35, 1-21, 1980).

Cualquier animal (ratón, conejo y similares), del que se sabe que produce anticuerpos, puede inmunizarse con el polipéptido elegido. Los métodos de inmunización son bien conocidos en la técnica. Dichos métodos incluyen la inyección subcutánea o intraperitoneal del polipéptido. Los expertos reconocerán fácilmente que la cantidad de polipéptido empleada para la inmunización variará en base al animal que se pretende inmunizar, la antigenicidad del polipéptido y el sitio de la inyección.

El polipéptido puede modificarse o administrarse en un adyuvante con el fin de aumentar la antigenicidad del péptido. Los métodos para aumentar la antigenicidad de un polipéptido son bien conocidos en la técnica. Tales procedimientos incluyen la unión de un antígeno con una proteína heteróloga (por ejemplo una globulina o β -galactosidasa) o la inclusión de un adyuvante durante la inmunización.

En lo que respecta a los anticuerpos monoclonales, se extraen células de bazo de animales inmunizados, se fusionan con células de mieloma y se convierten en células de hibridoma que produzcan anticuerpos monoclonales.

Puede aplicarse cualquiera de los muchos métodos bien conocidos de la técnica para identificar la célula de hibridoma que produce un anticuerpo de las características deseadas. Estos métodos incluyen la exploración de hibridomas con un ensayo ELISA, el análisis “western blot” o el radioinmunoensayo (Lutz y col., Exp. Cell. Res. 175, 109-124, 1988).

Los hibridomas que secretan los anticuerpos deseados se clonan y se determina el grupo o subgrupo aplicando procedimientos ya conocidos de la técnica (Campbell, Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology, lugar citado, 1984).

En los que respecta a los anticuerpos policlonales, se aísla el anticuerpo que contiene antisueros del animal inmunizado y se explora para determinar la presencia de anticuerpos que tengan la especificidad deseada aplicando uno de los procedimientos antes descritos.

En otra forma de ejecución que también se describe aquí, se marcan los anticuerpos antes descritos de un modo detectable. Se marcan los anticuerpos de un modo detectable con el uso de radioisótopos, marcadores de afinidad (por ejemplo la biotina, avidina y similares), marcadores enzimáticos (por ejemplo la peroxidasa de rábano rusticano, fosfatasa alcalina y similares), marcadores fluorescentes (por ejemplo el FITC o la rodamina y similares), átomos paramagnéticos y similares. Los procedimientos para realizar este marcado son bien conocidos en la técnica, véase, por ejemplo Sternberger y col., J. Histochem. Cytochem. 18, 315, 1970; Bayer y col., Meth. Enzym. 62, 308, 1979; Engval y col., Immunol. 109, 129, 1972; Goding, J. Immunol. Meth. 13, 215, 1976. Los anticuerpos marcados aquí descritos pueden utilizarse para ensayos “*in vitro*”, “*in vivo*” o “*in situ*” para identificar células o tejidos que expresen un péptido específico.

En otra forma de ejecución aquí descrita se inmovilizan los anticuerpos antes descritos sobre un soporte sólido. Los ejemplos de dichos soportes sólidos incluyen a los plásticos, por ejemplo el policarbonato, los hidratos de carbono complejos, por ejemplo la agarosa y la sefarsa, las resinas acrílicas, por ejemplo la poliácridamida, y las esferillas de látex. Los procedimientos para fijar dichos anticuerpos sobre el soporte sólido son bien conocidos en la técnica (Weir y col., “Handbook of Experimental Immunology”, 4ª ed., Blackwell Scientific Publications, Oxford, Inglaterra, capítulo 10, 1986; Jacoby y col., Meth. Enzym. 34, Academic Press, N.Y. 1974). Los anticuerpos inmovilizados aquí descritos pueden utilizarse para ensayos “*in vitro*”, “*in vivo*” e “*in situ*” así como para la inmunocromatografía.

Además, los expertos podrán adaptar fácilmente los procedimientos actualmente disponibles así como las técnicas, métodos y kits antes descritos en lo que respecta a los anticuerpos, para generar péptidos capaces de unirse a una secuencia peptídica específica con el fin de generar péptidos antipéptido de diseño racional, véase por ejemplo Hurby y col., “Application of Synthetic Peptides: Antisense Peptides”, en: Synthetic Peptides, A User’s Guide, W.H. Freeman, NY, pp. 289-307, 1992; y Kaspiczak y col., Biochemistry 28, 9230-8, 1989.

Los péptidos anti-péptido pueden generarse por una de dos maneras. En primer lugar, los péptidos anti-péptido pueden generarse reemplazando los restos aminoácidos básicos existentes en la secuencia peptídica del PCA3 por restos ácidos, pero manteniendo los grupos hidrófobos y polares sin carga. Por ejemplo los restos lisina, arginina, y/o histidina se reemplazan por restos ácido aspártico o ácido glutámico y los restos ácido glutámico se reemplazan por lisina, arginina o histidina.

VIII. *Un método para detectar un polipéptido de PCA3 o un anticuerpo en una muestra*

Se describe también un método para detectar un polipéptido de PCA3 en una muestra, que consiste en: a) poner en contacto la muestra con el anticuerpo (o proteína) antes descritos, en condiciones tales que se formen inmunocomplejos y b) detectar la presencia del anticuerpo fijado sobre el polipéptido. En detalle, los métodos consisten en incubar una muestra a analizar con uno o más anticuerpos aquí descritos y determinar si el anticuerpo se une a la muestra a analizar. Los niveles alterados de PCA3 en la muestra, por comparación con los niveles normales, indica que existe una enfermedad específica (p.ej. cáncer de próstata).

Se describe también un método para detectar un anticuerpo de PCA3 en una muestra, que consiste en: a) poner en contacto la muestra con la proteína de PCA3 antes descrita, en condiciones tales que se formen los inmunocomplejos y b) detectar la presencia de la proteína fijada sobre el anticuerpo o del anticuerpo fijado sobre la proteína. En detalle, los métodos consisten en incubar una muestra a analizar con una o más proteínas aquí descritas y determinar si el anticuerpo se fija sobre la muestra a analizar.

Las condiciones para incubar un anticuerpo con una muestra a analizar pueden variar. Las condiciones de incubación dependerán del formato empleado para el ensayo, de los métodos de detección empleados y del tipo y naturaleza del anticuerpo empleado para el ensayo.

Los expertos reconocerán fácilmente que se puede adaptar cualquiera de los formatos de ensayo inmunológico habitualmente disponibles (por ejemplo los radioinmunoensayos, ensayos inmunosorbentes asociados a enzimas, la técnica Ouchterlony basada en la difusión o el ensayo de cohete inmunofluorescente, para emplear los anticuerpos aquí descritos. Los ejemplos de tales ensayos se encontrarán en Chard, *An Introduction to Radioimmunoassay and Related Techniques*, Elsevier Science Publishers, Amsterdam, Holanda, 1986; Bullock y col., *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, Amsterdam, Holanda (1985).

Las muestras a analizar que pueden someterse al ensayo inmunológico aquí descrito incluyen a las células, proteínas o extractos de membranas de células, o líquidos biológicos como son la sangre, el suero, el plasma o la orina. La muestra a analizar empleada en el método antes descrito variará en función del formato del ensayo, de la naturaleza del método de detección y de los tejidos, células o extractos que se emplean como muestra a analizar. Los métodos para la preparación de extractos proteicos o extractos de membranas de células son bien conocidos en la técnica y pueden adaptarse fácilmente con el fin de obtener una muestra que sea susceptible de análisis con el sistema utilizado.

IX. *Un kit de diagnóstico que consta de la proteína o del anticuerpo de PCA3*

Se describe también un kit, que contiene todos los reactivos necesarios para llevar a cabo los métodos de detección recién descritos.

El kit puede contener: i) un primer recipiente que contiene el anticuerpo antes descrito, e ii) un segundo recipiente que contiene un conjugado formado por el reactivo de fijación sobre el anticuerpo y un marcador.

El kit puede contener: i) un primer recipiente que contiene la proteína antes descrita, y con preferencia ii) un segundo recipiente que contiene un conjugado formado por el reactivo de fijación sobre la proteína y un marcador. De modo más específico, un kit de diagnóstico contiene la proteína de PCA3 ya descrita antes para detectar anticuerpos del suero de animales o de seres humanos potencialmente infectados.

En otra forma preferida de ejecución, el kit contiene además lo siguiente: reactivos de lavado y reactivos capaces de detectar la presencia de anticuerpos fijados. Los ejemplos de reactivos de detección incluyen, pero no se limitan a: anticuerpos secundarios marcados, o como alternativa, si el anticuerpo primario está marcado, los reactivos cromóforos, enzimáticos o de fijación al anticuerpo, que son capaces de reaccionar con el anticuerpo marcado. El kit dividido en compartimentos puede ser del tipo antes descrito para los kits de sondas de ácido nucleico.

Los expertos reconocerán fácilmente que los anticuerpos aquí descritos pueden incorporarse sin problema a uno cualquiera de los formatos de kit actualmente disponibles, que son bien conocidos en la técnica.

X. Exploración de diagnóstico

Se da por supuesto que, aunque la discusión que sigue está centrada específicamente en los pacientes humanos, las enseñanzas pueden aplicarse también a cualquier animal que exprese el PCA3.

Los métodos de diagnóstico y exploración de la invención son especialmente útiles para un paciente del que se sospecha que corre el riesgo de desarrollar una enfermedad asociada con niveles de expresión alterados del PCA3 en fase a la historia familiar, o para un paciente en el que se desee diagnosticar la enfermedad relacionada con el PCA3 (p.ej. el cáncer de próstata).

Según la invención, la exploración presintomática de un individuo que necesite tal exploración es posible ahora mediante la codificación de DNA de la proteína del PCA3 o del gen de PCA3 de la invención o fragmentos del mismo. El método de exploración de la invención permite realizar un diagnóstico presintomático, incluido un diagnóstico prenatal, de la presencia de un gen de PCA3 ausente o aberrante en individuos y de este modo y formar una opinión sobre la probabilidad de que tal individuo desarrolle o haya desarrollado una enfermedad asociada con el PCA3. Esto es especialmente útil para la identificación de portadores que tengan el gen PCA3 alterado o ausente, por ejemplo individuos con una historia familiar de enfermedad asociada con el PCA3. El diagnóstico precoz es también deseable para maximizar la intervención temporal apropiada.

En una forma preferida de ejecución del método de exploración, se extrae una muestra de tejido de dicho individuo y se explora para determinar (1) la presencia del gen PCA3 "normal"; (2) la presencia del mRNA del PCA3 y/o (3) la presencia de la proteína del PCA3. El gen humano normal puede caracterizarse por ejemplo por detección de los modelos de digestión de restricción de los DNA "normales" frente a los DNA de pacientes, incluido el análisis por RFLP, empleando sondas de DNA preparadas contra la secuencia del PCA3 (o un fragmento funcional de la misma) acorde con la invención. De modo similar, el mRNA del PCA3 puede caracterizarse y compararse con el mRNA de PCA3 normal: (a) niveles y/o (b) tamaño existentes en la población humana que no corre riesgo de desarrollar una enfermedad asociada con el PCA3 empleando sondas similares. Finalmente, la proteína del PCA3 puede (a) detectarse y/o (b) cuantificarse empleando un ensayo biológico de la actividad del PCA3 o aplicando un ensayo inmunológico y anticuerpos de PCA3. Cuando se analiza la proteína del PCA3, es preferido el ensayo inmunológico porque es rápido. (1) Un modelo de tamaño de DNA de PCA3 aberrante y/o (2) tamaños o niveles de mRNA de PCA3 aberrantes y/o (3) niveles aberrantes de proteína de PCA3 indicarían que el paciente corre el riesgo de desarrollar la enfermedad asociada con el PCA3.

De modo más específico, aquí se proporciona un método de diagnóstico de la presencia o de la predisposición para desarrollar el cáncer de próstata en un paciente.

Los métodos de exploración y de diagnóstico de la invención no requieren que tenga que usarse para la sonda la secuencia entera que codifica al DNA del PCA3. Al contrario, solo es necesario emplear un fragmento o un tramo de ácido nucleico que sea suficiente para detectar la presencia del gen PCA3 en una preparación de DNA perteneciente a un individuo normal o afectado, la ausencia de tal gel o una propiedad física alterada de tal gen (por ejemplo un cambio en el modelo de migración electroforética). Con preferencia, puede utilizarse cualquiera de las sondas aquí descritas.

El diagnóstico prenatal puede efectuarse, si se desea, empleando cualquier método conocido de obtención de células fetales, incluida la amniocentesis, la extracción de muestras de la vellosidad corial (CVS) y la fetoscopia. Puede realizarse un análisis cromosómico prenatal para determinar si la porción del cromosoma que posee el gen PCA3 normal está presente en el estado heterocigótico.

XI. Tratamientos terapéuticos

A. Ácidos nucleicos terapéuticos

Un ácido nucleico terapéutico es un agente terapéutico que puede tener, pero no está limitado a ello, por lo menos uno de los siguientes efectos terapéuticos en una célula diana: inhibir la transcripción de una secuencia de DNA; inhibir la traducción de una secuencia de RNA; inhibir la transcripción inversa de una secuencia de RNA o de DNA; inhibir la modificación post-traducciona de una proteína; inducir la transcripción de una secuencia de DNA; inducir la traducción de una secuencia de RNA; inducir la transcripción inversa de una secuencia de RNA o de DNA; inducir una modificación post-traducciona de una proteína; transcripción del ácido nucleico como RNA; traducción del ácido nucleico como proteína o enzima; e incorporación del ácido nucleico al cromosoma de una célula diana para la expresión constitutiva o transitoria del ácido nucleico terapéutico.

Los efectos de los ácidos nucleicos terapéuticos pueden incluir, aunque no se imitan a ellos: desviar un gen defectuoso o el procesamiento de su expresión, por ejemplo en forma de RNA o DNA antisentido; inhibir la replicación vírica o la síntesis; terapia genética, por ejemplo la expresión de un ácido nucleico heterólogo que codifica a una proteína terapéutica o corregir una proteína defectuosa; modificar una expresión defectuosa o insuficiente de un RNA, por ejemplo en forma de hnRNA, mRNA, tRNA o rRNA; codificar un fármaco o profármaco o enzima que genera un compuesto, por ejemplo un fármaco o profármaco en células patológicas o normales que expresan al receptor quimérico; y todos los demás efectos terapéuticos conocidos.

En el método de tratamiento de una enfermedad asociada con el PCA3 (con preferencia, el cáncer de próstata) en un paciente que necesite tal tratamiento, puede proporcionarse un gen de PCA3 que no sea indicativo de un estado patológico a las células de dicho paciente de un modo y en una cantidad tales que permitan la expresión de la proteína de PCA3 proporcionada por dicho gen, durante un tiempo y en una cantidad suficiente para tratar a dicho paciente. Se aplica con preferencia la tecnología del reemplazo del gen ("knock out") que consiste en reemplazar un gen de PCA3 que provoca la enfermedad por un gen de PCA3 que no provoque la enfermedad (en concreto, el cáncer de próstata).

Se describen también composiciones farmacéuticas que contienen una cantidad eficaz por lo menos de un oligonucleótido antisentido de PCA3, en combinación con un vehículo farmacéuticamente aceptable. Estos oligonucleótidos antisentido incluyen, pero no se limitan a: por lo menos una secuencia de nucleótidos de una longitud de 12-500 bases que es complementaria de los exones 1, 2, 3, 4a-4d de PCA3; una secuencia de DNA de la SEQ ID NO: 1, 3, 4 ó 6; o una secuencia de DNA que codifique por lo menos a 4 aminoácidos de la SEQ ID NO: 2 ó SEQ ID NO: 7.

Como alternativa, el ácido nucleico de PCA3 puede combinarse con un vehículo lipófilo, por ejemplo uno de los muchos esteroides, incluido el colesterol, los colatos y el ácido desoxicólico. Un esteroide preferido es el colesterol.

Los ácidos nucleicos para la terapia con el gen de PCA3 y las composiciones farmacéuticas aquí descritas pueden administrarse por cualquier medio que permita alcanzar la finalidad pretendida. Por ejemplo, administración puede ser por vía parenteral, subcutánea, intravenosa, intramuscular, intraperitoneal o transdermal. La dosis administrada dependerá de la edad, la salud y el peso del receptor, del tipo de tratamiento simultáneo, si lo hubiera; de la frecuencia del tratamiento y de la naturaleza del efecto deseado.

Las composiciones aquí descritas incluyen a todas las composiciones que contienen un oligonucleótido antisentido de PCA3 en una cantidad eficaz para lograr una disminución de la expresión por lo menos de un gen de PCA3. Las necesidades concretas pueden variar, pero la determinación de los intervalos óptimos de la cantidad eficaz de cada componente podrán decidirla los expertos. Por ejemplo, el ácido nucleico de PCA3 puede administrarse a mamíferos, p.ej. humanos, en una dosis de 0,005 a 1 mg/kg/día, o una cantidad equivalente de una sal farmacéuticamente aceptable del mismo, por día, referido al peso corporal del mamífero a tratar.

Las formulaciones idóneas para la administración parenteral incluyen a las soluciones acuosas del ácido nucleico del PCA3 en una forma soluble en agua, por ejemplo sales solubles en agua. Además pueden administrarse suspensiones de los compuestos activos en suspensiones en aceite inyectables apropiadas. Los disolventes o vehículos lipófilos adecuados incluyen a los aceites grasos, por ejemplo el aceite de sésamo, o los ésteres de ácidos grasos sintéticos, por ejemplo el oleato de etilo o los triglicéridos. Las suspensiones acuosas inyectables pueden contener sustancias que aumenten la viscosidad de la suspensión e incluyen, por ejemplo, la carboximetilcelulosa sódica, la sorbita y/o el dextrano. La suspensión pueden contener opcionalmente estabilizantes.

En la técnica se conocen muchos sistemas de vectores para lograr la entrega a pacientes humanos que necesiten un gen o una proteína que falta en una célula. Pueden utilizarse, por ejemplo, sistemas de retrovirus, en especial sistemas de retrovirus modificados y en especial sistemas de virus de herpes simple (Gage y col., patente US-5,082,670). Estos métodos se describen, por ejemplo, en los artículos de Breakefield, X.A. y col., *The New Biologist* 3, 203-218, 1991; Huang, Q. y col., *Experimental Neurology* 115, 303-316, 1992; WO 93/03743; WO 90/0944; Taylor, WO 92/06693; Mulligan, R.C., *Science* 260, 926-932, 1993; y Brown y col., "Retroviral Vectors", en *DNA Cloning: A Practical Approach*, volumen 3, IRL Press, Washington, D.C., 1987. La entrega de una secuencia de DNA que codifique a una proteína de PCA3 expresada normalmente reemplazará con eficacia al gen PCA3 que provoca el estado patológico (p.ej. el cáncer de próstata).

Los medios para introducir el vector que lleva el ácido nucleico en una células incluyen, pero no se limitan a: la microinyección, la electroporación, la transducción o la transfección empleando DEAE-dextrano, la lipofección, el fosfato cálcico u otros procedimientos que los expertos ya conocen (*Molecular Cloning, A Laboratory Manual*, Sambrook y col., coordinadores, Cold Spring Harbor Press, Plainview, Nueva York, 1989).

En otra forma de ejecución de esta invención, un gen de PCA3 normal se expresa en forma de gen recombinante en una célula, de modo que las células pueden trasplantarse a un mamífero, con preferencia un ser humano que necesite la terapia genética. Para aplicar la terapia genética a un individuo se inserta en un vector una secuencia genética que codifica a la totalidad o una parte del gen de PCA3 y se introduce en una célula hospedante.

Pueden aplicarse otros métodos de terapia genética para transferir el ácido nucleico a un paciente, que se describe en: Chatterjee y 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 y col., *Seminars in Oncology* 23 (1), 159-171, 1996; y Dunbar, *Annu. Rev. Med.* 47, 11-20, 1996.

La especificidad de la expresión genética en células de cáncer de próstata puede conseguirse empleando secuencias reguladoras apropiadas, específicas de células, por ejemplo los intensificadores y promotores específicos de células.

Por lo tanto se puede emplear la terapia genética para aliviar la patología del PCA3 inhibiendo la expresión inapropiada de una forma particular del PCA3. Además, puede utilizarse la terapia genética para aliviar estas patología proporcionando un nivel de expresión apropiado de una forma concreta del PCA3. En este caso, pueden codificarse

las secuencias concretas del ácido nucleico del PCA3 con constructos de DNA o de RNA que se administran en forma de virus, del modo descrito antes.

B. Antagonistas y agonistas de PCA3

La capacidad de los antagonistas y agonistas de PCA3 para interferir o intensificar la actividad del PCA3 puede evaluarse en células que contengan el PCA3. Puede efectuarse un ensayo de la actividad de PCA3 en células para determinar la funcionalidad de la proteína de PCA3 en presencia de un agente que puede actuar como antagonista o agonista, y de este modo se identifican los agentes que interfieren o que intensifican la actividad del PCA3.

Los agentes explorados en los ensayos pueden ser, pero no se limitan a: anticuerpos, péptidos, hidratos de carbono, derivados de vitaminas u otros agentes farmacéuticos. Estos agentes pueden elegirse y explorarse 1) aleatoriamente, 2) mediante una selección racional o 3) por diseño, empleando por ejemplo técnicas de modelado de proteínas o ligandos (con preferencia, modelado computerizado).

En cuanto a la exploración aleatoria, se eligen los agentes, por ejemplo anticuerpos, péptidos, hidratos de carbono, agentes farmacéuticos y similares de modo aleatorio y se ensayan para determinar su capacidad de fijarse sobre o de estimular/bloquear la actividad de la proteína del PCA3.

Como alternativa, los agentes pueden elegirse o diseñarse de modo racional. Tal como se emplea aquí, la expresión "elegirse o diseñarse de modo racional" indica que el agente se elige en base a la configuración de la proteína del PCA3.

En una forma de ejecución, se describe un método de exploración de un antagonista o agonista que estimula o bloquea la actividad de PCA3, dicho método consiste en:

(a) incubar una célula que expresa al PCA3 con un agente a ensayar; y

(b) ensayar la célula para determinar la actividad de la proteína del PCA3 midiendo los efectos del agente en la fijación del PCA3 sobre el ATP.

En el ensayo anterior puede utilizarse cualquier célula, en el supuesto de que exprese una forma funcional del PCA3 y que pueda medirse la actividad de PCA3. Las células de expresión preferida son las células u organismos eucariotas. Dichas células pueden modificarse para que contengan secuencias de DNA que codifiquen al PCA3 aplicando procedimientos rutinarios ya conocidos en la técnica. Como alternativa, los expertos pueden introducir directamente en la célula al mRNA que codifica a la proteína del PCA3.

Empleando ligandos de PCA3 (ligandos que incluyen a los antagonistas y agonistas descritos antes) se describe también aquí un método para modular la actividad de la proteína del PCA3 en una célula. En general, los ligandos (antagonistas y agonistas), que se han identificado por estimular o bloquear la actividad de PCA3, pueden formularse de modo que el ligando pueda ponerse en contacto con una célula que expresa una proteína de PCA3 "*in vivo*". El contacto de tal célula con dicho ligando produce la modulación "*in vivo*" de la actividad de las proteínas del PCA3. En el supuesto de que no exista una barrera de formulación o una barrera de toxicidad, los ligandos identificados en los ensayos antes descritos serán eficaces para el uso "*in vivo*".

Se describe también un método de administración de PCA3 o de un ligando de PCA3 (incluidos los antagonistas y agonistas de PCA3) a un animal (con preferencia, un mamífero, en especial, un humano) en una cantidad suficiente para conseguir un nivel alterado de PCA3 en el animal. El PCA3 o ligando de PCA3 administrados podría conseguir específicamente funciones asociadas al PCA3. Además, dado que el PCA3 se expresa en células de cáncer de próstata, la administración de PCA3 o ligando de PCA3 podría utilizarse para alterar los niveles de PCA3 en dichas células.

Los expertos podrán apreciar que las cantidades a administrar para cualquier método concreto de tratamiento pueden determinarse sin problema. La dosis no debería ser muy amplia, para que no causara efectos secundarios adversos, por ejemplo reacciones cruzadas molestas, reacciones anafilácticas ni similares. En general, la dosis variará en función de la edad, el estado general de salud, el sexo y la extensión de la enfermedad en el paciente, las contraindicaciones, si las hubiera, y otras variables de este tipo, siendo necesario que el médico la ajuste oportunamente. La dosis puede variar entre 0,001 mg/kg y 50 mg/kg de PCA3 o de ligando de PCA3, en una o más administraciones al día, durante uno o varios días. El PCA3 o ligando de PCA3 puede administrarse por vía parenteral por inyección o por perfusión gradual a lo largo del tiempo. Puede administrarse por vía intravenosa, intraperitoneal, intramuscular o subcutánea.

Las preparaciones para la administración parenteral incluyen a las soluciones, suspensiones y emulsiones acuosas o no acuosas estériles. Los ejemplos de disolventes no acuosos son el propilenglicol, polietilenglicol, aceites vegetales, por ejemplo aceite de oliva y ésteres orgánicos inyectables, por ejemplo el oleato de etilo. Los vehículos acuosos incluyen al agua, soluciones, emulsiones o suspensiones alcohólico-acuosas, incluida la solución salina y los medios tamponados. Los vehículos parenterales incluyen la solución de cloruro sódico, la dextrosa de Ringer y el cloruro sódico, el lactato de Ringer y los aceites fijados. Los vehículos intravenosos incluyen a los rellenos líquidos y nutrientes, los rellenos de electrolito, por ejemplo los basados en la dextrosa de Ringer y similares. Los conservantes y otros adi-

tivos podrán estar también presentes, por ejemplo los agentes antimicrobianos, antioxidantes, agentes quelantes, gases inertes y similares. En general, véase el manual Remington's Pharmaceutical Science, 16ª ed., Mack Eds. (1980).

Se describe también una composición farmacéutica que contiene el PCA3 o un ligando de PCA3 en una cantidad suficiente para alterar la actividad asociada al PCA3 y un diluyente, vehículo o excipiente farmacéuticamente aceptable. Los expertos podrán determinar fácilmente las concentraciones apropiadas y las cantidades de las dosis unitarias del modo antes descrito (véase, por ejemplo, Remington's Pharmaceutical Sciences, 16ª ed., Osol, A., coord., editorial Mack, Easton, PA (1980) y WO 91/19008).

C. Inmunoterapia

Se describen también los anticuerpos de PCA3 mencionados anteriormente (con preferencia, anticuerpos murinos de PCA3 y anticuerpos quiméricos murino-humanos de PCA3 y fragmentos y regiones de los mismos), que inhiben o neutralizan la actividad biológica del PCA3 "*in vivo*" y son específicos del PCA3. Estos anticuerpos pueden utilizarse para finalidades terapéuticas en sujetos que sufren patología o estados patológicos asociados con la presencia de una expresión aberrante del PCA3. Los anticuerpos, fragmentos, regiones y derivados de los mismos, aquí descritos contienen con preferencia por lo menos una región que reconoce a un epítipo de PCA3 que tiene una actividad biológica inhibidora y/o neutralizadora "*in vivo*".

El tratamiento consiste en la administración parenteral de una dosis única o múltiple del anticuerpo, fragmento o derivado. Para el uso farmacéutico humano son preferidos los anticuerpos murino y quiméricos, que inhiben y/o neutralizan al PCA3, que poseen una gran actividad y potencia, así como sus fragmentos y regiones aquí descritos.

Los anticuerpos monoclonales aquí descritos pueden administrarse por cualquier medio que permita que el agente activo alcance el sitio de acción deseado en el cuerpo de un mamífero. Dado que las proteínas son susceptibles de digestión cuando se administran por vía oral, se empleará normalmente la administración parenteral, es decir, intravenosa, subcutánea, intramuscular, para optimizar la absorción.

Los anticuerpos monoclonales aquí descritos pueden administrarse en forma de agentes terapéuticos individuales o en forma de combinación con otros agentes terapéuticos. Pueden administrarse solos, pero por lo general se administran junto con un vehículo farmacéutico, elegido en función de la vía de administración y de la práctica farmacéutica estándar.

Como es obvio, la dosis administrada puede variar en función de factores conocidos, como son las características farmacodinámicas del agente concreto y su modo y vía de administración; la edad, el estado general de salud y el peso del receptor; la naturaleza y gravedad de los síntomas, el tipo de tratamiento concurrente, la frecuencia del tratamiento y el efecto deseado. Por lo general la dosis diaria de ingrediente activo se situará entre 0,1 y 100 miligramos por kilogramo de peso corporal. Para obtener los resultados deseados normalmente es eficaz una dosis de 0,5 a 50 y con preferencia de 1 a 10 miligramos por kilogramo al día, divididos en 1-6 dosis al día o en forma de entrega sostenida.

Las formas de dosificación (composiciones) idóneas para la administración interna se situarán por lo general entre 1 miligramos y 500 miligramos de ingrediente activo por unidad. En estas composiciones farmacéuticas, el ingrediente activo estará presente por lo general en una cantidad del 0,5 al 95% en peso, porcentaje referido al peso total de la composición.

Para la administración parenteral, el anticuerpo puede formularse en forma de solución, suspensión o emulsión o de polvo liofilizado en asociación con un vehículo parenteral farmacéuticamente aceptable. Los ejemplos de tales vehículos son el agua, solución salina, solución de Ringer, solución de dextrosa y albúmina de suero humano al 5%. Pueden utilizarse también liposomas y vehículos no acuosos, por ejemplo aceites fijados. El vehículo o polvo liofilizado pueden contener aditivos que mantengan la isotonía (p.ej. cloruro sódico, manita) y la estabilidad química (p.ej. tampones y conservantes). La formulación se esteriliza por los métodos habituales.

Los vehículos farmacéuticos idóneos se han descrito en la edición más reciente del manual Remington's Pharmaceutical Sciences, A. Osol, que es la obra de referencia en este ámbito.

Los anticuerpos murinos y quiméricos, sus fragmentos y regiones aquí descritos, sus fragmentos y derivados pueden utilizarse terapéuticamente en forma de inmunoconjugados (véase una revisión en Dillman, R.O., Ann. Int. Med. 111, 592-603, 1989). Pueden asociarse con proteínas citotóxicas, incluidas, pero sin limitarse a ellas: la ricina-A, la toxina de *Pseudomonas* y la toxina de la difteria. Las toxinas conjugadas con anticuerpos u otros ligandos ya son conocidas en la técnica (véase, por ejemplo, Olsnes, S. y col., Immunol. Today 10, 291-295, 1989). Las toxinas vegetales y bacterianas normalmente matan las células desbaratando el mecanismo de síntesis de las proteínas.

Los anticuerpos aquí descritos pueden conjugarse con otros tipos adicionales de restos terapéuticos que incluyen, pero no se limitan a: radionúclidos, agentes citotóxicos y fármacos. Los ejemplos de radionúclidos que pueden asociarse con los anticuerpos y administrarse "*in vivo*" a sitio de antígeno incluyen al ²¹²Bi, ¹³¹I, ¹⁸⁶Re y ⁹⁰Y, pero esta lista no pretende ser exhaustiva. Los radionúclidos despliegan su efecto citotóxico mediante la irradiación local de las células, que conduce a varias lesiones intracelulares, hecho ya conocido de la técnica de la radioterapia.

Los fármacos citotóxicos que pueden conjugarse con anticuerpos y después utilizarse para la terapia “*in vivo*” incluyen, pero no se limitan a: daunorrubicina, doxorubicina, metotrexato y mitomicina C. Los fármacos citotóxicos interfieren en los procesos celulares críticos, como son la síntesis del DNA, del RNA y de las proteínas. Una exposición más completa de estos grupos de fármacos, que ya son conocidos en la técnica, y de sus mecanismos de acción se encontrará p.ej. en Goodman, A.G. y col., Goodman y Gilman’s THE PHARMACOLOGICAL BASIS OF THERAPEUTICS, 7ª ed., Macmillan Publishing Co., 1985.

Los anticuerpos aquí descritos pueden utilizarse con ventaja en combinación con otros anticuerpos monoclonales o murinos y quiméricos, fragmentos y regiones de los mismos, o con linfoquinas o factores de crecimiento hematopoyético, etc., que sirven para aumentar el número o la actividad de las células efectoras, que interaccionan con los anticuerpos.

XII. Animales transgénicos no humanos de PCA3

Métodos para generar animales transgénicos no humanos

Los animales no humanos de la invención abarcan a todos los animales que tienen una interrupción o alteración transgénica del o de los genes endógenos (animales “knock-out”) y/o aquellos, en cuyo genoma se han introducido uno o más transgenes que dirigen la expresión del PCA3 humano. Es también preferida la introducción de ácidos nucleicos antisentido de PCA3.

Estos animales no humanos incluyen a los vertebrados, por ejemplo los roedores, los primates no humanos, las ovejas, el perro, la vaca, los anfibios, los reptiles, etc. Los animales no humanos preferidos se eligen entre las especies de mamíferos no humanos, con mayor preferencia entre el grupo de los roedores, incluidos los ratones, las ratas y con preferencia especial entre los ratones.

Los animales transgénicos de la invención son animales en los que se introducen por medios no naturales (es decir, por manipulación humana), uno o más genes que no existen en los animales de origen natural, p.ej. genes ajenos, genes endógenos manipulados genéticamente, etc. Los genes introducidos de modo no natural, también conocidos como transgenes, pueden proceder de la misma especie o de especies diferentes al animal, pero que no existen de modo natural en el animal en la configuración y/o lugar cromosómico conferido por el transgén. Los transgenes pueden contener secuencias de DNA foráneas, es decir, secuencias que normalmente no se hallan en el genoma del animal hospedante. Como alternativa o como adición, los transgenes pueden contener secuencias de DNA endógeno, que son anómalas por el hecho de que se han reordenado o se han mutado “*in vitro*” con el fin de alterar el modelo natural de expresión del gen “*in vivo*” o para alterar o eliminar la actividad biológica de un producto genético endógeno codificado por el gen (véase Watson, J.D. y col., en: Recombinant DNA, 2ª ed., W.H. Freeman & Co., Nueva 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 transgénicos no humanos de la invención se producen mediante la introducción de transgenes en la línea germinal del animal no humano. Se emplean células embrionarias diana de varios estadios de desarrollo para introducir los transgenes de la invención. Se aplican diferentes métodos en función del estadio de desarrollo de la o de las células embrionarias diana.

1. La microinyección de cigotos es el método preferido para la introducción de transgenes en los genomas animales en el curso de la puesta en práctica de la invención. Un cigoto, un huevo fecundado, que no haya sufrido la fusión pronuclear y la posterior división celular, es la célula diana preferida para la microinyección de secuencias de DNA transgénico. El pronúcleo de macho murino alcanza un tamaño de aproximadamente 20 micras de diámetro, una característica que permite la inyección reproducible de 1-2 picolitros de una solución que contiene las secuencias de DNA transgénico. El uso de un cigoto para la introducción de transgenes tiene la ventaja de que, en la mayoría de casos, las secuencias inyectadas de DNA transgénico se incorporarán al genoma del animal hospedante después de la primera división celular (Brinster y col., Proc. Natl. Acad. Sci. (USA) 82, 4438-4442, 1985). Como consecuencia de ello, todas las células de los animales transgénicos resultantes (animales fundadores) llevan incorporado de modo estable un transgén incorporado en un lugar genético concreto, que se denomina alelo transgénico. El alelo transgénico pone de manifiesto la herencia de Mendel: la mitad de la descendencia resultante del cruce de un animal transgénico con un animal no transgénico heredará el alelo transgénico, con arreglo a las leyes de Mendel de ordenamiento aleatorio.

2. Puede aplicarse también la integración vírica para introducir los transgenes de la invención en un animal. Los embriones en desarrollo se cultivan “*in vitro*” hasta el estadio de desarrollo conocido como blastocito. En aquel momento pueden infectarse los blastómeros con retrovirus apropiados (Jaenisch, R., Proc. Natl. Sci. (USA) 73, 1260-1264, 1976). La infección de los blastómeros se intensifica con la eliminación enzimática de la zona translúcida (Hogan y col., en: Manipulating the Mouse Embryo, Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 1986). Se introducen los transgenes con vectores víricos que son típicamente deficientes en replicación, pero que se mantienen competentes para la integración de secuencias de DNA asociadas con el virus, incluidas las secuencias de DNA transgénico unidas a tales secuencias víricas, en el genoma del animal hospedante (Jahner y col., Proc. Natl. Acad. Sci. (USA) 82, 6927-6931, 1985; Van der Putten y col., Proc. Natl. Acad. Sci. (USA) 82, 6148-6152, 1985). La transfección se consigue con facilidad y eficacia por cultivo de blastómeros en una monocapa de células que produzcan el vector vírico que contiene

al transgén (Van der Putten y col., Proc. Natl. Acad. Sci. (USA) 82, 6148-6152, 1985; Stewart y col., EMBO Journal 6, 383-388, 1987). Como alternativa, la infección puede realizarse en un estadio más tardío, por ejemplo el blastocelo (Jahner, D. y col., Nature 298, 623-628, 1982). En cualquier caso, la mayoría de animales fundadores transgénicos producidos por integración vírica serán mosaicos de alelos transgénicos; es decir, el transgén se incorpora solamente a un subconjunto de todas las células que forman el animal transgénico fundador. Además pueden suceder múltiples acontecimientos de integración vírica en un mismo animal fundador, generando múltiples alelos transgénicos, que se segregarán en futuras generaciones de la descendencia. La introducción de transgenes en las células de la línea germinal por este método es posible, pero probablemente tendrá lugar con una frecuencia baja (Jahner, D. y col., Nature 298, 623-628, 1982). Sin embargo, una vez se ha introducido por este método un transgén en las células de la línea germinal, es posible que surja una descendencia en la que el alelo transgénico esté presente en todas las células del animal, es decir, tanto en las células somáticas, como en las células de la línea germinal.

3. Las células del tronco embrionario (ES) pueden servir también de células diana para la introducción de los transgenes de la invención en animales. Las células ES se obtiene después de una preimplantación de embriones que se cultivan “*in vitro*” (Evans, M.J. y col., Nature 292, 154-156, 1981; Bradley, M.O. y col., Nature 309, 255-258, 1984; Gossler y col., Proc. Natl. Acad. Sci. (USA) 83, 9065-9069, 1986; Robertson y col., Nature 322, 445-448, 1986; Robertson, E.J., en Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, E.J., coord., IRL Press, Oxford, 1987, páginas 71-112). Las células ES, que son productos comerciales (suministrados p.ej. por Genome Systems, Inc., St. Louis, MO), pueden transformarse con uno o más transgenes por métodos ya establecidos (Lovell-Badge, R.H., en: Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, E.J., coord., IRL Press, Oxford, 1987, páginas 153-182). Las células ES transformadas pueden combinarse con un blastocito animal, después de lo cual las células ES colonizan el embrión y contribuyen a la línea germinal del animal resultante, que es una quimera (compuesta por 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., coord., IRL Press, Oxford, 1987, páginas 113-151). Una vez más, después de la introducción del transgén en las células de línea germinal por este método, puede producirse una descendencia en la que esté presente el alelo transgénico en todas las células del animal, es decir, tanto en las células somáticas como en las de la línea germinal.

Con independencia de su aparición, la introducción inicial de un transgén es un acontecimiento de Lamarck (y no de Mendel). Sin embargo, los transgenes de la invención pueden integrarse de forma estable en las células de la línea germinal y transmitirse a la descendencia del animal transgénico como lugares de Mendel. Otras técnicas transgénicas producen animales transgénicos de tipo mosaico, en los que algunas células llevan los transgenes y otras no. En los animales transgénicos de tipo mosaico, en los que las células de la línea germinal no llevan los transgenes, no tiene lugar la transmisión de los transgenes a la descendencia. No obstante, los animales transgénicos de tipo mosaico son capaces de presentar fenotipos asociados a los transgenes.

Los transgenes pueden introducirse en animales no humanos con el fin de proporcionar modelos animales de enfermedades humanas. Los transgenes que resultan de tales modelos animales incluyen, p.ej. a los transgenes que codifican a los productos genéticos mutantes asociados con un error innato del metabolismo de una enfermedad genética humana y transgenes que codifican a un factor humano requerido para conferir susceptibilidad (predisposición) para un patógeno humano (es decir, una bacteria, un virus u otro microorganismo patógeno) (Leder y col., patente US-5,175,383 (29 diciembre 1992); Kindt y col., patente US-5,183,949 (2 febrero 1993); Small y col., Cell 46, 13-18, 1986; Hooper y col., Nature 326, 292-295, 1987; Stacey y col., Nature 332, 131-136, 1988; Windle y col., Nature 343, 665-669, 1990; Katz y col., Cell 74, 1089-1100, 1993). Las mutaciones introducidas por métodos transgénicos abarcan a los alelos cero (“knock-out”), en los que una secuencia de DNA que codifica a un marcador seleccionable y/o detectable se ha sustituido por una secuencia genética normalmente endógena de un animal no humano. Los animales transgénicos no humanos resultantes que están predispuestos a una enfermedad, o en los que el transgén provoca una enfermedad, pueden utilizarse para identificar composiciones que inducen la enfermedad y para evaluar el potencia patogénico de las composiciones de las que se sabe o se sospecha que inducen la enfermedad (Berns, A.J.M., patente US-5,174,986 (29 diciembre 1992)), o para evaluar composiciones que puedan utilizarse para tratar la enfermedad o mejorar la evolución de sus síntomas (Scott y col., WO 94/12627, 1994).

La descendencia que ha heredado los transgenes de la invención puede distinguirse de los compañeros de camada que no han heredado los transgenes mediante el análisis del material genético de la descendencia determinando la presencia de biomoléculas que contienen secuencias únicas que corresponden a las secuencias de, o se han codificado por los transgenes de la invención. Por ejemplo, los líquidos biológicos que contienen polipéptidos codificados únicamente por el marcador seleccionable de los transgenes de la invención pueden someterse a inmunoensayo para determinar la presencia de los polipéptidos. Un método más sencillo y fiable de identificar la descendencia transgénica consiste en obtener una muestra de tejido de una extremidad del animal, p.ej. de la cola, y analizar la muestra para determinar la presencia de secuencias de ácido nucleico que corresponden a la secuencia de DNA de de una porción única o de porciones de los transgenes de la invención, por ejemplo el marcador seleccionable de las mismas. La presencia de estas secuencias de ácido nucleico puede determinarse p.ej. por análisis de hibridación (“Southern”) con secuencias de DNA que corresponden a porciones únicas del transgén, por análisis de los productos de las reacciones PCR empleando secuencias de DNA de una muestra como sustratos y oligonucleótidos derivados de la secuencia de DNA del transgén, etc.

La presente invención se describe a continuación con mayor detalle mediante los siguientes ejemplos no limitantes.

Ejemplo 1

Aislamiento y caracterización del cDNA y del DNA genómico del PCA3

5 Con el fin de identificar nuevos marcadores del cáncer de próstata, se realiza un análisis de visualización diferencial (Liang y col., Science 257, 967-971, 1992) para identificar los genes sobreexpresados en carcinomas prostáticos en comparación con la próstata normal; se extraen el RNA total de tejido prostático de hiperplasia benigna y maligna de los mismos pacientes. Se emplean veinte combinaciones diferentes de cebadores (cuatro cebadores anclados, cinco cebadores arbitrarios), de ellos se identifican once que aparentemente expresan los mRNA de modo diferencial (es decir, se sobreexpresan de modo consistente en todos los carcinomas estudiados y no se expresan en el tejido normal o de BPH). Se emplean los fragmentos de DNA complementario (cDNA) como sondas para el análisis Northern Blot para confirmar la sobreexpresión consistente en los tumores prostáticos empleados para la visualización diferencial. Una de las sondas (llamada DD3, un cDNA de 486 bp) detecta dos transcritos principales (2,3 y 4,0 kb) que están muy sobreexpresados en 47 de los 50 tumores prostáticos humanos estudiados, mientras que no se observa expresión (o niveles muy bajos de la misma) de estos transcritos en el tejido normal o de BPH de los mismos pacientes.

Para obtener un clon de cDNA de longitud completa se construye una biblioteca de cDNA empleando un mRNA aislado de tejido de tumor prostático primario humano. Se obtienen 250 clones DD3 positivos por exploración de esta biblioteca. Se purifican 80 clones y se determina la secuencia de nucleótidos de estos clones por análisis de secuenciación automática.

Se explora una biblioteca genómica construida con DNA genómico de placenta humana, clonado en λ FIX2, empleando como sonda el DD3. Se obtienen cuatro clones diferentes, dos de ellos situados hacia el extremo 5' del gen (λ FIX-ME3 y -ME4) y dos clones situados hacia el extremo 3' del gen (λ FIX-ME1 y -ME2). Se subclona el extremo 5' de λ FIX-ME4 y se emplea como sonda para explorar la biblioteca genómica. Se aíslan tres clones únicos nuevos (λ FIX-IH1, IH2 y IH6).

De los 80 clones de cDNA analizados, se observa que están presentes cuatro transcritos diferentes, debido al empalme alternativo y a la poliadenilación alternativa. Los análisis de secuencia de los clones genómicos, si se comparan con los clones de cDNA, pone de manifiesto la estructura genómica del gen PCA3. Están presentes tres intrones y 4 exones. El primer intrón tiene una longitud aproximada de 20 kb.

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

La segunda especie de cDNA, encontrada en aproximadamente el 15% de los cDNA clonados, contiene los exones 1, 3, 4a, 4b y 4c, y surge como empalme alternativo del segundo exón (no presente en este cDNA) y termina en una señal diferente de poli-A-adición (consenso real) (figura 1).

La tercera especie de cDNA contiene los exones 1, 3, 4a y 4b y es la que se encuentra con mayor frecuencia (aproximadamente en el 65% de los 80 clones) (figura 1). Este cDNA es con gran probabilidad el causante del transcrito más importante observado con el análisis Northern blot (2 kb).

La cuarta especie de cDNA detectada contiene los exones 1, 3 y 4a, que representan aprox. el 15% de clones y termina después del 4a, que es el sitio de paro del clon DD3 original (figura 1). La señal de poli-A-adición aquí presente está muy próxima a la secuencia de consenso.

El PCA3 es un gen, en el que tiene lugar un empalme alternativo significativo (así como poliadenilación alternativa), como ponen de manifiesto los transcritos de diferentes tamaños observados en el análisis Northern Blot y se identifican diferentes tipos de clones. Tal como se ha mencionado antes, pueden identificarse otras variantes de empalme, ya que virtualmente es posible cualquier combinación de exones. Por ejemplo, recientemente se ha identificado un clon de cDNA que tiene los exones 2, 3, 4a, 4b y 4c. Por tanto, parece que se han aislado clones que representan virtualmente todas las combinaciones posibles de exones.

Se ha identificado una de estas variantes de empalme por secuenciación de un clon llamado λ DD3.6. El λ DD3.6 es un clon λ gt11 identificado y aislado por exploración de una biblioteca de cDNA obtenida a partir de RNA de próstata de un varón de 25 años (facilitada por Clonetech) con una sonda PCA3. El λ DD3.6 contiene el exón 3, 4a, 4b, 4c y 4d. Sin embargo, este clon de cDNA contiene también secuencias de intrón (parte del intrón 2, también el intrón 3).

En la figura 3 se presenta una comparación de los clones depositados PMB9 y λ DD3.6.

Se examinan las diferentes combinaciones de exones por análisis computerizado para identificar los marcos de lectura abiertos (ORF) y para predecir la región que codifica a la proteína. El ORF más largo es también la región que con mayor probabilidad codifica a la proteína.

El ORF más largo, de 153 nucleótidos, codifica a un péptido pequeño de 51 aminoácidos, el PCA3. El PCA3 es codificado por una parte del exón 3 y 4a. El pequeño tamaño de la proteína sugiere que la proteína actúa con

gran probabilidad como molécula mensajera y tiene el potencial de ser secretada por las células. La secuencia de nucleótidos de los exones 1-4a-d y la secuencia de aminoácidos del PCA3 se representan en la fig. 2 y fig. 5 (SEQ ID NO: 1 y 6 y 2 y 7, respectivamente).

Los expertos reconocerán fácilmente que el clon de cDNA que contiene la secuencia de ácido nucleico presentada en la SEQ ID NO: 6 y representada en la fig. 5 puede obtenerse del modo antes descrito por aislamiento y caracterización de los clones de cDNA del PCA3. Por ejemplo, y como ya es sabido en la técnica, pueden utilizarse las sondas específicas por lo menos de uno de los exones 1, 2, 3, 4a, 4b, 4c y 4d del extremo 5' para aumentar la probabilidad de tener un clon de cDNA de PCA3 de longitud completa. Pueden utilizarse por ejemplo placas de 96 hoyos, por ejemplo, para explorar un gran número de clones de cDNA positivos de PCA3, empleando las sondas antes mencionadas. Obviamente, los clones positivos de PCA3 pueden también secuenciarse, hecho ya conocido y descrito aquí, hasta que se obtenga un clon de cDNA deseado.

Es también posible obtener un clon de cDNA que contenga la secuencia representada en la SEQ ID NO: 6 y en la fig. 5, empleando cebadores específicos de PCA3 y un método de amplificación, por ejemplo la PCR. Por ejemplo, la tecnología PCR con cebadores específicos de los últimos extremos 5' y 3' del cDNA del PCA3 podría utilizarse para amplificar un producto deseado (casi 4 kb) a partir del RNA aislado por ejemplo de tumores prostáticos y clonar los productos de la PCR. Sin embargo, dado que la amplificación PCR puede introducir errores, lo más probable es que sea necesaria la secuenciación del cDNA completo.

Los expertos ya conocen que un clon de cDNA que tenga la secuencia representada en la SEQ ID NO: 6 y en la fig. 5 puede construirse también empleando los clones aquí descritos (o clones recién aislados) y los métodos convencionales de ingeniería genética.

Por ejemplo, puede construirse un clon de cDNA de longitud completa de este tipo empleando los clones depositados pMB9 y λ DD3.6. Un ejemplo no limitante de una estrategia para construir un clon de cDNA de este tipo consiste en la secuencia de ácido nucleico de la SEQ ID NO: 6 y fig. 5, que se describe a continuación.

Se digiere el DNA fago de λ DD3.6 por completo con NdeI, y se aísla el fragmento NdeI aproximadamente de 2 kb en un gel de agarosa. Este fragmento contiene parte del exón 4b del PCA3, los exones 4c + 4d y aprox. 50 nucleótidos del DNA fago. Los extremos de este fragmento de 2 kb se rellenan con la DNA-polimerasa del fragmento de Klenow y con dNTP, el fragmento de extremo romo se liga a continuación con los sitios HindII/SmaI del Bluescript SK. La pérdida del sitio HindII del Bluescript por digestión del HincII y SmaI es esencial para los siguientes pasos de clonación en esta estrategia particular (ver más abajo). Cabe reseñar que los sitios NdeI están también presentes en el fago λ gt11, dando lugar a varios fragmentos adicionales, algunos de los cuales están muy próximos a 2 kb (es decir, un fragmento de 1,8 kb y un fragmento de 2,5 kb). Sin embargo es muy sencillo separar estas bandas diferentes en un gel de agarosa. La orientación correcta de la inserción del fragmento NdeI de 2 kb de extremo romo del λ DD3.6 en el Bluescript (llamado constructo PCA3-X) puede verificarse por digestión simple con SacI, que proporcionaría un fragmento de -0,45 y -4,5 kb por tinción con bromuro de etidio en los geles de agarosa. El análisis de la secuencia puede realizarse para confirmar la identidad del inserto PCA3.

Después puede digerirse el constructo PCA3-X por completo con HindIII y BamHI y puede aislarse en el gel de agarosa un inserto de vector de 4,8 kb. Esto provoca la eliminación de -0,2 kb de DNA del inserto. El pMB9 puede digerirse simultáneamente por completo con BamHI y HindIII y el inserto de 1,9 kb (que contiene los exones de PCA3 1, 2, 3, 4a y la mayor parte del exón 4b) aislado en un gel de agarosa. El inserto derivado del pMB9 se liga al sitio BamHI/HindIII del constructo PCA3-X. El constructo resultante, el PCA3-Y, contiene el cDNA completo del PCA3, excepto los 22 primeros nucleótidos del exón 1 (véase más abajo y la fig. 4). Estos 22 nucleótidos pueden añadirse al cDNA del PCA3 por corte del constructo PCA3-Y y el oligo-(74)-mer (SEQ ID NO: 8) por completo con BamHI y PstI y ligando el oligómero al constructo PCA3-Y, de este modo se obtiene el constructo PCA3-Z. El análisis de la secuencia de nucleótidos puede realizarse para verificar que el oligonucleótido se ha ligado correctamente (es decir, para confirmar que se ha ligado precisamente un oligonucleótido y no toda la serie de oligonucleótidos). Obviamente, la secuenciación del cDNA resultante del PCA3-Z podrá realizarse para verificar la integridad de la secuencia de ácido nucleico.

La exploración de un papel híbrido de células somáticas pone de manifiesto que el gen que codifica al PCA3 está localizado en el cromosoma humano 9. Empleando una mezcla de cuatro clones genómicos afines al PCA3 como sonda para hibridar con los cromosomas de metafase de los linfocitos humanos, se mapea el PCA3 en 9q21-22 (véase también la figura 1).

La conservación del gen PCA3 durante la evolución se ha estudiado por análisis Southern Blot y ha puesto de manifiesto que un homólogo de este gen está presente en el mono, la vaca, el caballo, la oveja, la cabra y el cerdo. El gen está también presente en el perro y en el gato. Por comparación, el gen que codifica el PSA solamente se ha hallado en los primates.

Ejemplo 2

Expresión de PCA3 específica de la próstata

Después de desarrollar cebadores específicos del PCA3 se realiza un análisis por RT-PCR empleando el RNA de diversos tejidos humanos normales. Con 40 ciclos de PCR se amplifican productos relacionados con el PCA3 en próstata normal y en tejidos de BPH. La expresión del PCA3 es muy específica de la próstata, ya que no se ha podido amplificar ningún producto de PCA3 en estas condiciones en los siguientes tejidos humanos normales: arterias, cerebro, mama, vejiga, colon, duodeno, corazón, hígado, pulmón, ovarios, páncreas, placenta, vesículas seminales, musculatura esquelética, piel, médula espinal, bazo y testículos. Tampoco han podido detectarse en las líneas celulares de cáncer de próstata humano ALVA-31, DU145, JCA-1, PPC-1, PC3 y TSU-Pr1 ningún producto de PCR relacionado con el PCA3. En la línea celular LNCaP se ha podido obtener un producto después de 40 ciclos de PCR (mientras que en las mismas condiciones puede obtenerse un producto en tumores prostáticos después de 20 ciclos). La tecnología empleada para evaluar la expresión del PCA3 específica de la próstata puede adaptarse a un ensayo de diagnóstico para el cáncer de próstata. Puede adaptarse además para la identificación del origen prostático de una metástasis.

Se ha realizado además un análisis RT-PCR semicuantitativo para comparar la expresión del PCA3 con la del PSA (antígeno específico de la próstata) y del PSM (antígeno específico de la membrana prostática) y para determinar si el análisis PCA3-RT-PCR puede utilizarse para distinguir entre muestras prostáticas malignas y benignas. Después de cuantificar la reacción RT se emplean 10 ng de cDNA para la reacción PCR y, como control, se examina también la beta-2-microglobulina. Los productos de PCA3 encontrados permiten una distinción clara entre las muestras benignas y malignas en 23 de los 25 casos estudiados, mientras que el PSA y el PSM no consiguen hacer esta distinción: se encuentran cantidades aproximadamente iguales de producto en las muestras normales y en las tumorales. La expresión del PSA y del PCA3 se compara también por análisis Northern Blot, que indica claramente la mayor especificidad tumoral del PCA3. Se observa una sobreexpresión del PCA3 en carcinomas prostáticos 20 veces mayor que la encontrada en tejidos normales o tejidos de BPH. Esta es claramente diferente de la expresión del PSM y del PSA, ambos se expresan menos en tejidos malignos que en tejidos benignos. Por tanto, parece que el PCA3 es un buen marcador para el diagnóstico del cáncer de próstata.

Un marcador tumoral ideal del cáncer de próstata debería ser capaz no solo de distinguir positivamente entre tejidos benignos y malignos, sino también ser capaz de predecir el desenlace clínico (curación o progresión) de pacientes afectados por esta enfermedad. Los datos indican que el nivel de expresión del PCA3 tiende a guardar una correlación claramente positiva con el grado del tumor.

Se emplea el RISH (e eventualmente la inmunohistoquímica) para determinar si existe correlación o no entre la sobreexpresión del PCA3, el grado del tumor, el estadio y el desenlace clínico. Se dispone de un seguimiento clínico de larga duración tanto para las muestras sumergidas en parafina como para las congeladas. Aplicando un análisis por la imagen asistido con ordenador se realiza una cuantificación de los niveles de expresión del PCA3 detectado por el RISH y se normalizan frente a una referencia externa (Tamimi y col., Cancer Res. 53, 5512-16, 1993; Tamimi y col., B. J. Cancer, 1996). Se realizará un análisis de regresión multivariable incluido el Gleason: grado, estadio tumoral patológico, estadio tumoral clínico, niveles de PSA y expresión del PCA3 para determinar si el PCA3 puede predecir con precisión la progresión y si tiene un valor (adicional) para el pronóstico.

Los ensayos de reacción en cadena de la polimerasa con la transcriptasa inversa (RT-PCR) se han desarrollado para detectar células micrometastásicas hematogénicas ocultas que, de otro modo, pasarían desapercibidas con las modalidades de clasificación tumoral (estadiaje) actualmente disponibles. Estos ensayos de RT-PCR se han realizado ya en pacientes de cáncer de próstata y otras enfermedades malignas. Puede utilizarse un ensayo RT-PCR muy sensible (encajado) (u otros tipos de amplificación, incluyendo pero sin limitarse a ellos: NASBA, PCR, QB rep., SOA, TMA y LCR (Winn-Deen, J. Clin., Liquid Assay 19, 21-26, 1996) para detectar las células del cáncer de próstata en la sangre circulante de pacientes de cáncer de próstata y para identificar los pacientes que corren el riesgo de tener o desarrollar metástasis. Los ensayos incluirán los controles apropiados (p.ej. β -2-microglobulina) y se realizarán de modo semicuantitativo (es decir, se cuantificará la síntesis del cDNA y se emplearán cantidades iguales para alimentar los análisis por PCR).

Se realizarán estudios de estadiaje molecular en el contexto más amplio del programa BIOMED II (marcadores del cáncer de próstata). En este amplio estudio realizado en colaboración se estudiarán el PSA y el PSM así como otros marcadores potencialmente interesantes del cáncer de próstata. Se están recogiendo ya muestras de sangre de pacientes, a los que se ha diagnosticado la enfermedad prostática en las instituciones que participan. Se ha iniciado la optimización de la recogida y de la manipulación de las muestras de sangre de los pacientes para la detección de células tumorales circulantes. Por ejemplo, el uso de tubo del tipo vacutainer™ CPT (Beckton-Dickinson) para la recogida de la sangre y la purificación de los leucocitos de la sangre periférica en combinación con un procedimiento de extracción de RNA del tipo Trizol™ (basado en el tiocianato de guanidinio) permiten la preparación de RNA idónea en calidad y cantidad para los análisis de PCR. El uso de cebadores específicos del PCA3 para amplificar los transcritos de PCA3 en el RNA extraído de la sangre de pacientes de cáncer de próstata pone de manifiesto que se ha podido detectar la presencia de células de cáncer de próstata en la sangre circulante, no solo en el caso de pacientes con metástasis probadas, sino también en el caso de pacientes de enfermedad supuestamente localizada. Se realizarán estudios más amplios en una población más numerosa de pacientes y se buscará la correlación entre los datos clínicos y el seguimiento para determinar el valor predictivo del PCA3 para el paciente individual de cáncer de próstata.

ES 2 324 503 T3

El análisis RT-PCR encajado (nested) (o métodos similares de amplificación) debería ser decisivo para determinar si hay órganos (todavía no estudiados) que expresan el PCA3. Deberán estudiarse, por ejemplo, la glándula de Cowper (que tiene el mismo origen embrionario que la próstata) y también la glándula de Skene (“homólogo” femenino de la próstata) por su posible reacción frente al PCA3.

En una muestra de tejido prostático “normal”, que contenía un 10% de células tumorales, se ha detectado la expresión del PCA3, lo cual indica que el PCA3 tiene una gran sensibilidad como marcador tumoral. De este modo, la expresión del PCA3 se ha detectado también en unas pocas muestras de BPH, en las que después se ha encontrado que contienen pequeñas zonas de células tumorales. El nivel de expresión del PCA3 en cánceres de próstata indica una tendencia hacia una correlación positiva con el grado del tumor. Estos datos se basan en el análisis de autorradiografía resultantes de la hibridación Northern Blot.

La observación de que la expresión del PCA3 parece aumentar con la pérdida de diferenciación es distinta de los que se ha publicado del PSA, ya que los niveles de expresión del PSA disminuyen con la pérdida de diferenciación (Hakalahti y col., Int. J. Cancer 55, 590-597, 1993). Hay una sobreexpresión del PCA3 en carcinomas prostáticos que es por lo menos 20 veces mayor que la registrada en tejidos normales o de BPH. Esta es claramente diferente de la expresión del PSA, de la que se ha publicado que disminuye en los tejidos malignos frente a los benignos. La expresión del PCA3 se ha detectado en 4 de las 4 metástasis estudiadas.

Ejemplo 3

Identificación de un sitio de inicio de transcripción de PCA3

Con el fin de determinar el sitio de inicio de la transcripción del PCA3 se han realizado análisis de extensión con cebador, mapeo con nucleasa S1 y 5'RACE (rapid amplification of cDNA ends). Se ha encontrado que el principal sitio de inicio de la transcripción está situado dentro de un intervalo de 4 nucleótidos (figura 4).

Los resultados de estos ensayos siguen ampliando el tamaño del cDNA en dirección 5' con otros 22 nt con respecto a la secuencia de cDNA del pMB9 (SEQ ID NO: 1 y figura 2). Esta secuencia adicional de polinucleótido 5' se representa también en la SEQ ID NO: 6 y figura 5).

REIVINDICACIONES

1. Una molécula de ácido nucleico asociada con el cáncer de próstata y derivadas de las secuencias de exón de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), que contiene una secuencia de polinucleótido elegida entre el grupo formado por:

(a) una secuencia de nucleótidos del exón 1 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 1-98 de la SEQ ID NO: 1;

(b) una secuencia de nucleótidos del exón 2 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 99-263 de la SEQ ID NO: 1;

(c) una secuencia de nucleótidos del exón 3 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 264-446 de la SEQ ID NO: 1;

(d) una secuencia de nucleótidos del exón 4a de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 447-985 de la SEQ ID NO: 1;

(e) una secuencia de nucleótidos del exón 4b de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 986-2037 de la SEQ ID NO: 1;

(f) una secuencia de nucleótidos que contiene secuencias de combinaciones de exones de las secuencias de nucleótidos de (a), (b), (c), (d) y (e);

(g) una secuencia de nucleótidos del exón 1 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 1-120 de la SEQ ID NO: 6;

(h) una secuencia de nucleótidos del exón 2 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 121-285 de la SEQ ID NO: 6;

(i) una secuencia de nucleótidos del exón 3 de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 286-468 de la SEQ ID NO: 6;

(j) una secuencia de nucleótidos del exón 4a de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 469-1007 de la SEQ ID NO: 6;

(k) una secuencia de nucleótidos del exón 4b de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 1008-2066 de la SEQ ID NO: 6;

(l) una secuencia de nucleótidos del exón 4c de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 2067-2622 de la SEQ ID NO: 6;

(m) una secuencia de nucleótidos del exón 4d de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), existente entre los nucleótidos 2623-3582 de la SEQ ID NO: 6;

(n) una secuencia de nucleótidos que contiene secuencias de combinaciones de exones de las secuencias de nucleótidos de (g), (h), (i), (j), (k), (l) y (m);

(o) una secuencia de nucleótidos que es completamente complementaria de una cualquiera de las secuencias de nucleótidos de (a) a (n); y

(p) una secuencia de nucleótidos que se hibrida específicamente en condiciones de gran restricción a 68°C con una cualquiera de las secuencias de nucleótidos de (a) a (o), dicha secuencia de nucleótidos consta de 10-1000 nucleótidos.

2. La molécula de ácido nucleico de la reivindicación 1, en la que dicha secuencia de nucleótidos de (p) está formada por:

(a) de 10 a 50 nucleótidos;

(b) de 10 a 35 nucleótidos; o

(c) de 20 a 35 nucleótidos.

3. La molécula de ácido nucleico de reivindicación 1, en la que dicha secuencia de nucleótidos de (p) se hibrida con una secuencia definida en la SEQ ID NO: 1.

ES 2 324 503 T3

4. La molécula de ácido nucleico de reivindicación 1, en la que dicha secuencia de nucleótidos de (f) o (n) se hibrida una secuencia definida en la SEQ ID NO: 6.
5. Un método para detectar una molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3) en una muestra, que consiste en:
- (a) poner en contacto dicha muestra con la molécula de ácido nucleico según una cualquiera de las reivindicaciones 1-4; y
 - (b) detectar la presencia de dicha molécula fijada sobre la molécula de ácido nucleico de PCA3, empleando una secuencia de nucleótidos derivada de dicha molécula de ácido nucleico del antígeno de cáncer de próstata (PCA3).
6. El método de reivindicación 5, en el que dicha molécula de ácido nucleico de PCA3 contiene los nucleótidos 1-98 de la SEQ ID NO: 1 o los nucleótidos 1-120 de la SEQ ID NO: 6.
7. El método de reivindicación 5 ó 6, en el que dicha molécula de ácido nucleico de PCA3 contiene los nucleótidos 264-446 de la SEQ ID NO: 1 o los nucleótidos 286-468 de la SEQ ID NO: 6.
8. El método de reivindicación 5, 6 ó 7, en el que dicha secuencia en (b) está definida en la SEQ ID NO: 4.
9. El método de una cualquiera de las reivindicaciones 5-8, en el que dicha detección consiste en una detección de un radiomarcador, de fluorescencia, de quimioluminiscencia o de cualquier combinación de ellas.
10. Una combinación de cebadores para la amplificación de una secuencia de exón de la molécula de ácido nucleico del antígeno del cáncer de próstata (PCA3), en la que dichos cebadores se hibridan específicamente con la molécula de ácido nucleico de la reivindicación 1.
11. La combinación de cebadores de la reivindicación 10, en la que cada uno de los cebadores está formado por 10-50 nucleótidos.
12. El método de una cualquiera de las reivindicaciones 5-9 que consiste en:
- (a) amplificar dicha molécula de ácido nucleico de PCA3 en dicha muestra, que consiste en poner en contacto dicha muestra con dicha combinación de cebadores según la reivindicación 10 ú 11; y
 - (b) detectar la presencia de dicha molécula de ácido nucleico de PCA3 visualizando un producto de la amplificación de (a).
13. El método de la reivindicación 12, en el que dicha muestra se elige entre el grupo formado por:
- (a) una muestra que contiene tejido o células de próstata;
 - (b) el RNA de tejido o células humanas;
 - (c) el DNA de tejido o células humanas; y
 - (d) orina, sangre o suero humanos.
14. El método de la reivindicación 13, en el que dicha muestra es orina humana.
15. Una molécula de ácido nucleico recombinante que contiene una secuencia de vector y el ácido nucleico. molécula según una cualquiera de las reivindicaciones 1-4.
16. La molécula de ácido nucleico recombinante de la reivindicación 15, que contiene un promotor eficaz para iniciar la transcripción de dicha molécula de ácido nucleico en una célula hospedante.
17. Una célula que contiene la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1-4, o la molécula de ácido nucleico recombinante de la reivindicación 15 ó 16.
18. Un organismo no humano que contiene la molécula de ácido nucleico recombinante según la reivindicación 15 ó 16.
19. La combinación de cebadores de la reivindicación 10, en la que dichos cebadores se hibridan con secuencias de combinaciones de exones de dicha molécula de ácido nucleico del antígeno de cáncer de próstata (PCA3).

ES 2 324 503 T3

20. La molécula de ácido nucleico de una cualquiera de las reivindicaciones 1-4, el método de una cualquiera de las reivindicaciones 5-9, 12, 13 y 14, o la combinación de cebadores de la reivindicación 10, 11 ó 19, dicha molécula de ácido nucleico o dichos cebadores están formados por lo menos por 15 nucleótidos.

5 21. La molécula de ácido nucleico, el método o la combinación de cebadores de la reivindicación 20, dicha molécula de ácido nucleico o dichos cebadores tienen un tamaño elegido entre grupo formado por:

(a) 20-35 nucleótidos;

10 (b) 20-50 nucleótidos; y

(c) 20-100 nucleótidos.

15 22. La molécula de ácido nucleico, el método o la combinación de cebadores de la reivindicación 21, dicha molécula de ácido nucleico o dichos cebadores están formados por 20-35 nucleótidos.

20

25

30

35

40

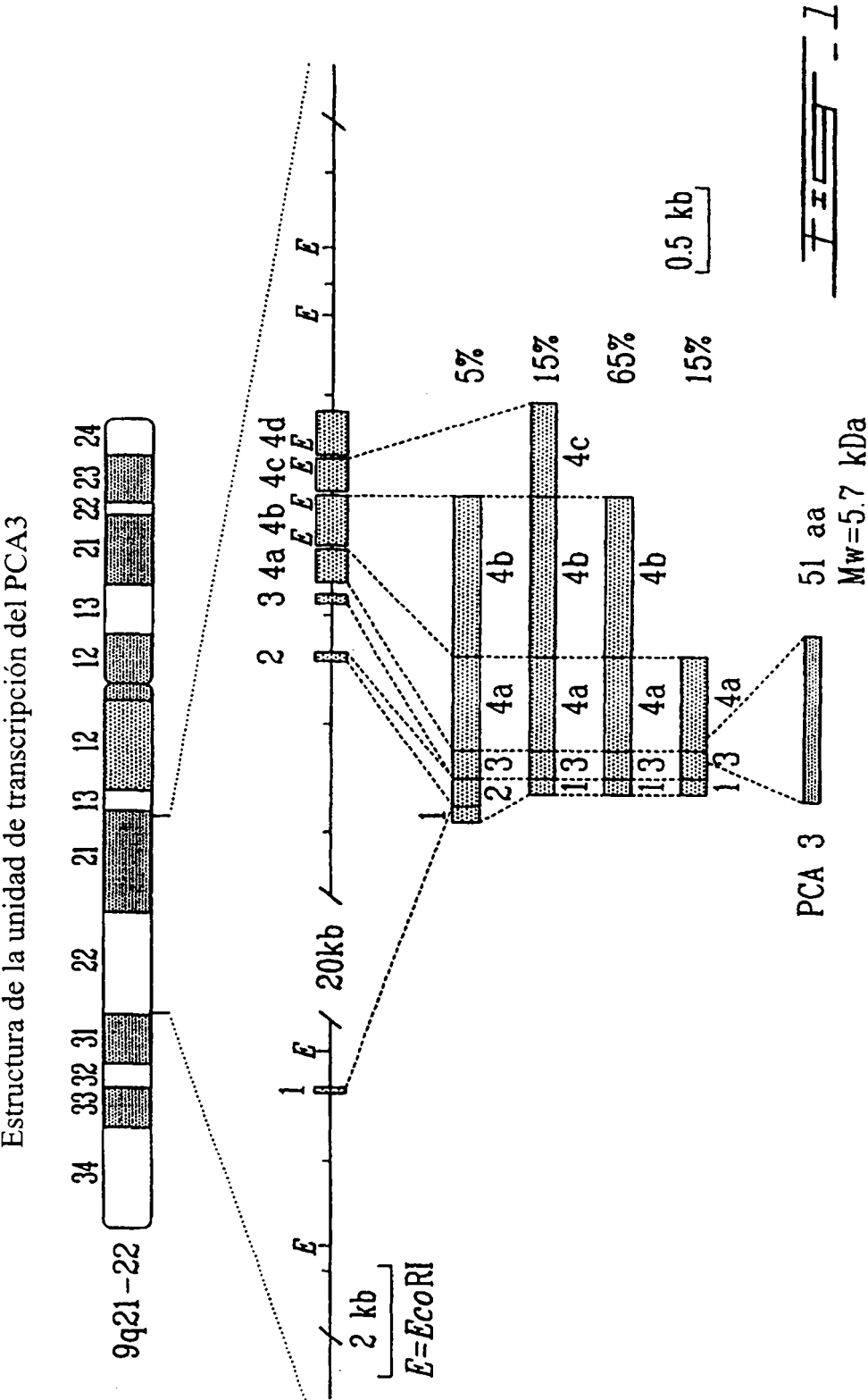
45

50

55

60

65



```

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

Fig. 2A

1) EXON1, 2) EXON2, 3) EXON3, 4) EXON4A, 5) EXON4b, PCA3 (aminoácidos)
(las señales de poli-A-adición están subrayadas)
CON) : Range= 1 to 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
                                     Fig. 2B
```

CON 61 GAAGATCTGC ATGGTGGGAA GGACCTGATG ATACAGAGGA ATTACAACAC ATATACTTAG

2 23 tgtttcaatg aacaccaaga taaataagtg aagagtagt ccgctgtgag tctcctcagt

CON 121 TGTTTCAATG AACACCAAGA TAAATAAGTG AAGAGCTAGT CCGCTGTGAG TCTCCTCAGT

2 83 gacacagggc tggatcacca tcgacggcac tttctgagta ctcagtgag caaagaaaag

CON 181 GACACAGGGC TGGATCACCA TCGACGGCAC TTTCTGAGTA CTCAGTGCAG CAAAGAAAAG

2 142 actacagaca tctcaatggc aggg

3 1 gtgaga aataagaaaag gctgctgact ttaccatctg

CON 240 ACTACAGACA TCTCAATGGC AGGGTGAGA AATAAGAAAG GCTGCTGACT TTACCATCTG

3 37 aggccacaca tctgctgaaa tggagataat taacatcact agaaacagca agatgacaat

CON 300 AGGCCACACA TCTGCTGAAA TGGAGATAAT TAACATCACT AGAAACAGCA AGATGACAAT

7181-2C

3 97 ataatgtcta agtagtgac atgttttttg cacatttcc agcccccttt aaatatcca cacaca
 CON 360 ATAATGTCTA AGTAGTGAC ATGTTTTTG CACATTTC AGCCCCCTTT AAATATCCA CACACA
 PCA3 1 M F L H I S S P F K Y P H T
 3 158 caggaagca caaaaggaa gcacagag
 4 1 a tccctggga gaaatgccc ggccgccat cttg99
 CON 421 CAGGAAGCA CAAAAGGAA GCACAGAGA TCCCTGGGA GAAATGCCC GGCCGCCAT CTTGGG
 PCA3 15 Q E A Q K E A Q R S L G E M P G R H L G
 4 35 tcacgcatg agcctcgcc ctgtgcctg gtccccgctt gtgAGGGAA GGACATTAG AAA
 CON 481 TCATCGATG AGCCTCGCC CTGTGCCCTG GTCCCGCTT GTGAGGAA GGACATTAG AAA
 PCA3 35 S S M S L A L C L V P L V R E G H ***
 4 93 ATGAATTGAT GTGTTCCCTTA AAGGATGGC AGGAAAACAG ATCCTGTTGT GGATATTAT
 CON 538 ATGAATTGAT GTGTTCCCTTA AAGGATGGC AGGAAAACAG ATCCTGTTGT GGATATTAT

7151-20

4 153 TTGAACGGGA TTACAGATT TAAATGAAGT CACAAAGTGA GCATTACCAA TGAGAGGAAA
 CON 598 TTGAACGGGA TTACAGATT TAAATGAAGT 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 CATATTCTA ACCCTCAAAA
 CON 778 GCCCTGCTGC CTAAACTGTG CGTTCATAAC CAAATCATTT CATATTCTA ACCCTCAAAA
FE-2E

4 393 CAAAGCTGTT GTAATATCTG ATCTCTACGG TTCCTTCTGG GCCCAACATT CTCATATAT
 CON 838 CAAAGCTGTT GTAATATCTG ATCTCTACGG TTCCTTCTGG GCCCAACATT CTCATATAT
 4 453 CCAGCCACAC TCATTTTAA TATTAGTTC CCAGATCTGT ACTGTGACCT TTCTACACTG
 CON 898 CCAGCCACAC TCATTTTAA TATTAGTTC CCAGATCTGT ACTGTGACCT TTCTACACTG
 4 513 TAGAATAACA TTA~~CT~~CATT~~T~~ TGTTC~~AAA~~
 5 1 GA CCTTCGTGT TGCTGCCTAA TATGTAGCTG
 CON 958 TAGAATAACA TTA~~CT~~CATT~~T~~ TGTTC~~AAA~~GA CCTTCGTGT TGCTGCCTAA TATGTAGCTG
 5 33 ACTGTTTTTC CTAAGGAGTG TTCTGGCCCA GGGATCTGT GAACAGGCTG GGAAGCATCT
 CON 1018 ACTGTTTTTC CTAAGGAGTG TTCTGGCCCA GGGATCTGT GAACAGGCTG GGAAGCATCT
 5 93 CAAGATCTTT CCAGGGTTAT ACTTACTAGC ACACAGCATG ATCATTACGG AGTGAATTAT
FF-2F

CON 1078 CAAGATCTTT CCAGGGTTAT ACTTACTAGC ACACAGCATG ATCATTACGG AGTGAATTAT

5 153 CTAATCAACA TCATCCTCAG TGTCTTTGCC CATACTGAAA TTCAATTCCC ACTTTTGTC

CON 1138 CTAATCAACA TCATCCTCAG TGTCTTTGCC CATACTGAAA TTCAATTCCC ACTTTTGTC

5 213 CCATTCTCAA GACCTCAAAA TGTCAATCCA TTAATATCAC AGGATTAACT TTTTITTTTA

CON 1198 CCATTCTCAA GACCTCAAAA TGTCAATCCA TTAATATCAC AGGATTAACT TTTTITTTTA

5 273 ACCTGGAAGA ATTCAATGTT ACATGCAGCT ATGGGAATT AATTACATAT TTTGTTTTCC

CON 1258 ACCTGGAAGA ATTCAATGTT ACATGCAGCT ATGGGAATT AATTACATAT TTTGTTTTCC

5 333 AGTGCAAAGA TGACTAAGTC CTTTATCCCT CCCCTTTGTT TGATTTTTTT TCCAGTATAA

CON 1318 AGTGCAAAGA TGACTAAGTC 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 CTAACCTTGTA
 CON 1438 AGCCTTATCT GTCATCACCA TCAACCCCTC CCATNYSACC TAAACAAAAT CTAACCTTGTA

5 513 ATTCCCTTGAA CATGTCAGN CATACTRT TCCTTCTGCC TGAGAAGCTC TTCCTTGTCT
 CON 1498 ATTCCCTTGAA CATGTCAGN CATACTRT TCCTTCTGCC TGAGAAGCTC TTCCTTGTCT

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

FE-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 CTTCATTTCT CTATCTCTAT CACAATATCC AACAAAGCTTT TCACAGAATT CATGCAGTGC
 CON 1798 CTTCATTTCT CTATCTCTAT CACAATATCC AACAAAGCTTT TCACAGAATT CATGCAGTGC
 5 873 AAATCCCCAA AGGTAACCTT TATCCATTTC ATGGTGAGTG CGCTTTAGAA TTTTGGCAAA

FEI - 21

CON 1858 AAATCCCCAA AGGTAACTT TATCCATTTC ATGGTGAGTG CGCTTTAGAA TTTTGGCAAA

5 933 TCATACTGGT CACTTATCTC AACTTTGAGA TGTGTTTGTC CTGTAGTTA ATTGAAAGAA

CON 1918 TCATACTGGT CACTTATCTC AACTTTGAGA TGTGTTTGTC CTGTAGTTA ATTGAAAGAA

5 993 ATAGGGCACT CTTgtgagcc actttagggt tcactcctgg caataaaaa ttacaaaga

CON 1978 ATAGGGCACT CTTgtgagcc actttagggt tcactcctgg caataaaaa ttacaaaga

FiS - 2J

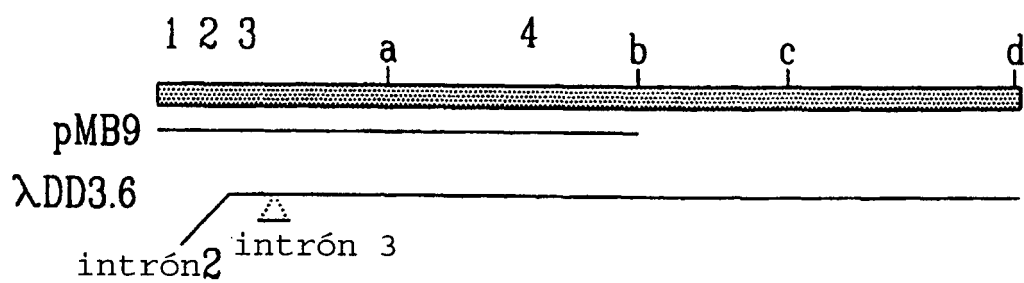
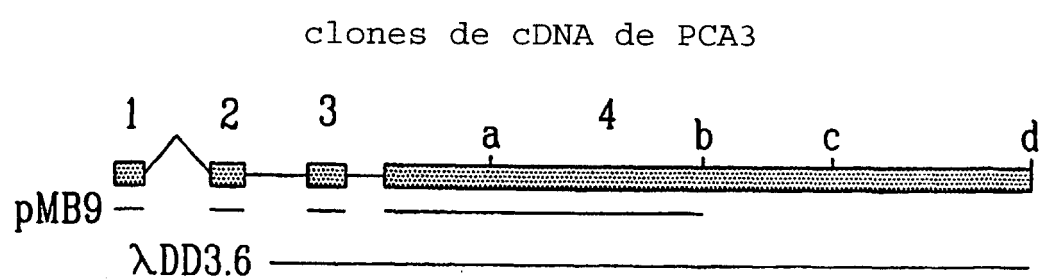


Fig. 3

^{5'} PE | ^{5'} RACE ^{5'} | ^{5'} RACE cDNAs
 ▼ ▼ ▼ ▼ ▼
 ...g ACAGAAG AAATAGCAAG TGCCGAGAAG CTGGCATCAG
 ↳ TSS ↳ pMB9
 AAAAACAGAG GGGAGATTG TGTGGCTGCA GCCGAGGGAG ACCAGGAAGA
 TCTGCATGGT GGAAGGACC TGATGATACA GAG gt ...
 sitio de empalme

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
FEF-5A

```

Secuencia del cDNA de PCA3 y de la proteína del PCA3

```

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

```

361 TAGAAACAGC AAGATGACAA TATAATGTCT AAGTAGTGAC **ATG**TTTTTGC **ATG**TTTTTGC ACATTTCAG
PCA3 1 M F L H I S S

421 CCCCTTTTAA TATCCACACA CACAGGAAGC ACAAAAGGAA GCACAGAGAT CCCTGGGAGA
PCA3 8 F F K Y P H T Q E A Q K E A Q R S L G E

481 AATGCCCCGC CGCCATCTTG GGTCAATCGAT GAGCCTCGCC CTGTGCCCTGG TCCCGCTTGT
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 GGAAACACAGA
PCA3 48 R E G H *

601 TCCTGTTGTG GATATTTATT TGAACGGGAT TACAGATTG AAATGAAGTC ACAAAGTGAG

661 CATTACCAAT GAGAGGAAAA CAGACGAGAA AATCTTGATG GCTTCACAAG ACATGCAACA

721 AACAAAATGG AATACTGTGA TGACATGAGG CAGCCCAAGCT GGGGAGGAGA TAACCACGGG

781 GCAGAGGGTC AGGATTCTGG CCTGCTGCC TAAACTGTGC GTTCATAACC AAATCATTC

841 ATATTCTAA CCCTCAAAAC AAAGCTGTTG TAATATCTGA TCTCTACGGT TCCTTCTGGG

901 CCCAACATTC TCCATATATC CAGCCACACT CATTTTAAAT ATTTAGTTCC CAGATCTGTA

961 CTGTGACCTT TCTACACTGT AGAATAACAT TACTCATTTT GTTCAAAGAC CCTTCGTGTT

1021 GCTGCCCTAAT ATGTAGCTGA CTGTTTTTCC TAAGGAGTGT TCTGGCCCAG GGGATCTGTG

725-51

1081 AACAGGCTGG GAAGCATCTC AAGATCTTTC CAGGGTTATA CTTACTAGCA CACAGCATGA
 1141 TCATTACGGA GTGAATTATC TAATCAACAT CATCCTCAGT GTCTTTGCCC ATACTGAAAT
 1201 TCATTCCCA CTTTGTGCC CATTCCTAAG ACCTCAAAAT GTCATTCCAT TAATATCACA
 1261 GGATTAACTT TTTTTTTTAA CCTGGAAGAA TTCAATGTTA CATGCAGCTA TGGGAATTTA
 1321 ATTACATATT TTGTTTTTCCA GTGCAAGAT 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 TCCTTGCTC TTAAATCTAG AATGATGTAA AGTTTGAAT AAGTTGACTA
 1621 TCTTACTTCA TGCAAAGAAG GGACACATAT GAGATTCATC ATCACATGAG ACAGCAAATA
 1681 CTAAAAGTGT AATTTGATTA TAAGAGTTTA GATAAATATA TGAAATGCAA GAGCCACAGA
 1741 GGAATGTTT ATGGGCACG TTTGTAAGCC TGGGATGTGA AGCAAAGGCA GGAACCTCA
 1801 TAGTATCTTA TATAATATAC TTCATTCTC TATCTCTATC ACAATATCCA ACAAGCTTTT
 1861 CACAGAATTC ATGCAGTGCA AATCCCCAAA GGTAACCTTT ATCCATTTC TGGTGAGTGC

ES - 50

1921 GCTTTAGAAT TTTGGCAAAT CATACTGGTC ACTTATCTCA ACTTTGAGAT GTGTTTGTCC
 1981 TTGTAGTTAA TTGAAAGAAA TAGGGCACTC TTGTGAGCCA CTTAGGGTT CACTCCTGGC
 2041 AATAAAGAAT TTACAAAGAG CTACTCAGGA CCAGTTGTTA AGAGCTCTGT GTGTGTGTGT
 2101 GTGTGTGTGT GAGTGATACAT GCCAAAGTGT GCCTCTCTCT CTTGACCCCAT TATTTCAGAC
 2161 TTA AAAACAAG CATGTTTTCA AATGGCACTA TGAGCTGCCA ATGATGTATC ACCACCATAT
 2221 CTCATTATTC TCCAGTAAAT GTGATAATAA TGTCACTCTGT TAACATAAAA AAAGTTTGAC
 2281 TTCACAAAAG CAGCTGGAAA TGGACAACCA CAATATGCAT AAATCTAACT CCTACCATCA
 2341 GCTACACACT GCTTGACATA TATTGTTAGA AGCACCTCGC ATTTGTGGGT TCTCTTAAGC
 2401 AAAATACTTG CATTAGGTCT CAGCTGGGGC TGTGCATCAG GCGGTTTGAG AAATATTCAA
 2461 TTCTCAGCAG AAGCCAGAAT TTGAATTCCC TCATCTTTTA GGAATCATTT ACCAGGTTTG
 2521 GAGAGGATTC AGACAGCTCA GGTGCTTTCA CTAATGTCTC TGAACCTCTG TCCCTCTTTG
 2581 TGTTTCATGGA TAGTCCAATA AATAATGTTA TCTTTGAACT GATGCTCATA GGAGAGAATA
 2641 TAAGAACTCT GAGTGATATC AACATTAGGG ATTCAAAGAA ATATTAGATT TAAGCTCACA
 2701 CTGGTCAAAA GGAACCAAGA TACAAAGAAC TCTGAGCTGT CATCGTCCCC ATCTCTGTGA

FE-5E

2761 GCCACAACCA ACAGCAGGAC CCAACGCATG TCTGAGATCC TTAAATCAAG GAAACCAGTG
 2821 TCATGAGTTG AATTCTCCTA TTATGGATGC TAGCTTCTGG CCATCTCTGG CTCTCCTCTT
 2881 GACACATATT AGCTTCTAGC CTTTGCTTCC ACGACTTTTA TCTTTTCTCC AACACATCGC
 2941 TTACCAATCC TCTCTCTGCT CTGTTGCTTT GGACTTCCCC ACAAGAATT CAACGACTCT
 3001 CAAGTCTTTT CTTCCATCCC CACCACTAAC CTGAATTGCC TAGACCCTTA TTTTATTAA
 3061 TTTCCAATAG ATGCTGCCTA TGGGCTAATA TTGCTTTAGA TGAACATTAG ATATTTAAAG
 3121 TCTAAGAGGT TCAAAATCCA ACTCATTATC TTCTCTTTCT TTCACCTCCC CTGCTCCTCT
 3181 CCTATATTA CTGATTGACT GAACAGGATG GTCCCCAAGA TGCCAGTCAA ATGAGAAACC
 3241 CAGTGGCTCC TTGTGGATCA TGCATGCAAG ACTGCTGAAG CCAGAGGATG ACTGATTACG
 3301 CCTCATGGGT GGAGGGGACC ACTCCTGGGC CTTCGTGATT GTCAGGAGCA AGACCTGAGA
 3361 TGCTCCCCTGC CTTCAGTGTC CTCTGCATCT CCCCTTTCTA ATGAAGATCC ATAGAATTG
 3421 CTACATTGA GAATTCCAAT TAGGAACTCA CATGTTTAT CTGCCCCATC AATTTTTAA
 3481 ACTTGCTGAA AATTAAGTTT TTTCAAAATC TGTCCTTGTA AATTACTTTT TCTTACAGTG
 3541 TCTTGGCATA CTATATCAAC TTTGATTCTT TGTACAACT TT

TSF - 5F

ES 2 324 503 T3

LISTA DE SECUENCIAS

(1) INFORMACIÓN GENERAL:

(i) SOLICITANTE:

(A) NOMBRE: BUSSEMAKERS, Marion J.G.

(B) CALLE: St Geertruidestraat 27

(C) CIUDAD: 6521 KG Nijmegen

(E) PAÍS: Holanda

(F) CÓDIGO POSTAL (ZIP): 6521 KG

(ii) TÍTULO DE LA INVENCION: PCA3, GENES PCA3 y MÉTODOS DE USO

(iii) NÚMERO DE SECUENCIAS: 8

(iv) FORMA LEÍBLE CON ORDENADOR:

(A) TIPO DE MEDIO: disquete (floppy disk)

(B) ORDENADOR: IBM PC compatible

(C) SISTEMA OPERATIVO: PC-DOS/MS-DOS

(D) PROGRAMA INFORMÁTICO: PatentIn Release #1.0, Version #1.30 (EPO)

(vi) DATOS DE SOLICITUD ANTERIOR:

(A) NÚMERO DE SOLICITUD: US 60/041,836

(B) FECHA DE DEPÓSITO: 10-ABRIL-1997

(2) INFORMACIÓN DE LA SEQ ID NO: 1:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 2037 pares de bases

(B) TIPO: ácido nucleico

(C) NÚMERO DE HEBRAS: doble

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: cDNA

(ix) CARACTERÍSTICAS:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 379..531

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 1:

```
AGAAGCTGGC ATCAGAAAAA CAGAGGGGAG ATTTGTGTGG CTGCAGCCGA GGGAGACCAG      60
GAAGATCTGC ATGGTGGGAA GGACCTGATG ATACAGAGGA ATTACAACAC ATATACTTAG      120
TGTTTCAATG AACACCAAGA TAAATAAGTG AAGAGCTAGT CCGCTGTGAG TCTCCTCAGT      180
```

ES 2 324 503 T3

	GACACAGGGC TGGATCACCA TCGACGGCAC TTTCTGAGTA CTCAGTGCAG CAAAGAAAGA	240
	CTACAGACAT CTCAATGGCA GGGGTGAGAA ATAAGAAAGG CTGCTGACTT TACCATCTGA	300
5	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	
10	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	
15	ATG CCC GGC CGC CAT CTT GGG TCA TCG 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	
20	GTC CCG CTT GTG AGG GAA GGA CAT TAGAAAATGA ATTGATGTGT TCCTTAAAGG	561
	Val Pro Leu Val Arg Glu Gly His	
	45 50	
	ATGGGCAGGA AAACAGATCC TGTGTGGAT ATTTATTTGA ACGGGATTAC AGATTTGAAA	621
25	TGAAGTCACA AAGTGAGCAT TACCAATGAG AGGAAAACAG ACGAGAAAAT CTTGATGGCT	681
	TCACAAGACA TGCAACAAAC AAAATGGAAT ACTGTGATGA CATGAGGCAG CCAAGCTGGG	741
30	GAGGAGATAA CCACGGGGCA GAGGGTCAGG ATTCTGGCCC TGCTGCCTAA ACTGTGCGTT	801
	CATAACCAAA TCATTTCATA TTTCTAACCC TCAAAACAAA GCTGTTGTAA TATCTGATCT	861
	CTACGGTTCC TTCTGGGCCC AACATTCTCC ATATATCCAG CCACACTCAT TTTTAATATT	921
35	TAGTTCCCAG ATCTGTACTG TGACCTTTCT ACACTGTAGA ATAACATTAC TCATTTTGTT	981
	CAAAGACCCT TCGTGTGCT GCCTAATATG TAGCTGACTG TTTTTCCTAA GGAGTGTCT	1041
40	GGCCCAGGGG ATCTGTGAAC AGGCTGGGAA GCATCTCAAG ATCTTTCCAG GGTTATACTT	1101
	ACTAGCACAC AGCATGATCA TTACGGAGTG AATTATCTAA TCAACATCAT CCTCAGTGTC	1161
	TTTGCCCATTA CTGAAATTCA TTTCCCACTT TTGTGCCCAT TCTCAAGACC TCAAAATGTC	1221
45	ATTCCATTAA TATCACAGGA TTAACTTTTT TTTTAACTT GGAAGAATTC AATGTTACAT	1281
	GCAGCTATGG GAATTAAATT ACATATTTTG TTTTCCAGTG CAAAGATGAC TAAGTCCTTT	1341
50	ATCCCTCCCC TTTGTTTGAT TTTTTCCTCA GTATAAGTT AAAATGCTTA GCCTTGTA	1401

ES 2 324 503 T3

```

GAGGCTGTAT ACAGCACAGC CTCTCCCAT CCCTCCAGCC TTATCTGTCA TCACCATCAA 1461
CCCCCTCCCAT NYSACCTAAA CAAAATCTAA CTTGTAATTC CTTGAACATG TCAGGNCATA 1521
5 CATTRTTCCT TCTGCCTGAG AAGCTCTTCC TTGTCTCTTA ANTCTAGAAT GATGTAAAGT 1581
TTTGAATAAG TTGACTATCT TACTTCATGC AAAGAAGGGA CACATATGAG ATTCATCATC 1641
10 ACATGAGACA GCAAATACTA AAAGTGTAAT TTGATTATAA GAGTTTAGAT AAATATATGA 1701
AATGCAAGAK CCACAGAGGG AATGTTTATG GGGCACGTTT GTAAGCCTGG GATGTGAAGM 1761
AAAGGCAGGG AACCTCATAG TATCTTATAT AATATACTTC ATTTCTCTAT CTCTATCACA 1821
15 ATATCCAACA AGCTTTTCAC AGAATTCATG CAGTGCAAAT CCCCAAAGGT AACCTTTATC 1881
CATTTTCATGG TGAGTGCCTT TTAGAATTTT GGCAAATCAT ACTGGTCACT TATCTCAACT 1941
20 TTGAGATGTG TTTGTCCTTG TAGTTAATTG AAAGAAATAG GGCACCTCTG TGAGCCACTT 2001
TAGGGTTCAC TCCTGGCAAT AAAGAATTTA CAAAGA 2037

```

(2) INFORMACIÓN DE LA SEQ ID NO: 2:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 51 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 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

```

(2) INFORMACIÓN DE LA SEQ ID NO: 3:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1872 pares de bases

(B) TIPO: ácido nucleico

(C) NÚMERO DE HEBRAS: doble

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: cDNA

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 3:

ES 2 324 503 T3

	AGAAGCTGGC ATCAGAAAAA CAGAGGGGAG ATTTGTGTGG CTGCAGCCGA GGGAGACCAG	60
	GAAGATCTGC ATGGTGGGAA GGACCTGATG ATACAGAGGT GAGAAATAAG AAAGGCTGCT	120
5	GACTTTACCA TCTGAGGCCA CACATCTGCT GAAATGGAGA TAATTAACAT CACTAGAAAC	180
	AGCAAGATGA CAATATAATG TCTAAGTAGT GACATGTTTT TGCACATTTT CAGCCCCTTT	240
10	AAATATCCAC ACACACAGGA AGCACAAAAG GAAGCACAGA GATCCCTGGG AGAAATGCCC	300
	GGCCGCCATC TTGGGTCATC GATGAGCCTC GCCCTGTGCC TGGTCCCGCT TGTGAGGGAA	360
	GGACATTAGA AAATGAATTG ATGTGTTCTT TAAAGGATGG GCAGGAAAAC AGATCCTGTT	420
15	GTGGATATTT ATTTGAACGG GATTACAGAT TTGAAATGAA GTCACAAAGT GAGCATTACC	480
	AATGAGAGGA AAACAGACGA GAAAATCTTG ATGGCTTCAC AAGACATGCA ACAAACAAAA	540
20	TGGAATACTG TGATGACATG AGGCAGCCAA GCTGGGGAGG AGATAACCAC GGGGCAGAGG	600
	GTCAGGATTC TGGCCCTGCT GCCTAAACTG TGCCTTCATA ACCAAATCAT TTCATATTTT	660
	TAACCCCTCAA AACAAAGCTG TTGTAATATC TGATCTCTAC GGTTCCTTCT GGGCCCAACA	720
25	TTCTCCATAT ATCCAGCCAC ACTCATTTTT AATATTTAGT TCCCAGATCT GTACTGTGAC	780
	CTTTCTACAC TGTAGAATAA CATTACTCAT TTTGTTCAAA GACCCCTTCGT GTTGCTGCCT	840
30	AATATGTAGC TGAATGTTTT TCCTAAGGAG TGTTCGGCC CAGGGGATCT GTGAACAGGC	900
	TGGGAAGCAT CTCAAGATCT TTCCAGGGTT ATACTTACTA GCACACAGCA TGATCATTAC	960
35	GGAGTGAATT ATCTAATCAA CATCATCCTC AGTGTCTTTG CCCATACTGA AATTCATTTT	1020
	CCACTTTTGT GCCCATTTCTC AAGACCTCAA AATGTCATTC CATTAAATATC ACAGGATTAA	1080
	CTTTTTTTTT TAACCTGGAA GAATTCAATG TTACATGCAG CTATGGGAAT TTAATTACAT	1140
40	ATTTTGTTTT CCAGTGCAAA GATGACTAAG TCCTTTATCC CTCCCCTTTG TTTGATTTTT	1200
	TTTCCAGTAT AAAGTTAAAA TGCTTAGCCT TGTACTGAGG CTGTATACAG CACAGCCTCT	1260
45	CCCCATCCCT CCAGCCTTAT CTGTCATCAC CATCAACCCC TCCCATNYS A CCTAAACAAA	1320
	ATCTAACTTG TAATTCCTTG AACATGTCAG GNCATACATT RTTCCTTCTG CCTGAGAAGC	1380
	TCTTCCTTGT CTCTTAANTC TAGAATGATG TAAAGTTTTG AATAAGTTGA CTATCTTACT	1440
50	TCATGCAAAG AAGGGACACA TATGAGATTC ATCATCACAT GAGACAGCAA ATACTAAAAG	1500
	TGTAATTTGA TTATAAGAGT TTAGATAAAT ATATGAAATG CAAGAKCCAC AGAGGGAATG	1560
55	TTTATGGGGC ACGTTTGTA GCTGGGATG TGAAGMAAAG GCAGGGAACC TCATAGTATC	1620
	TTATATAATA TACTTCATTT CTCTATCTCT ATCACAATAT CCAACAAGCT TTTCACAGAA	1680
60	TTCATGCAGT GCAAATCCCC AAAGGTAACC TTTATCCATT TCATGGTGAG TGCGCTTTAG	1740
	AATTTTGGCA AATCATACTG GTCACCTATC TCAACTTTGA GATGTGTTTG TCCTTGTAGT	1800
	TAATTGAAAG AAATAGGGCA CTCTTGTGAG CCACCTTAGG GTTCACTCCT GGCAATAAAG	1860
65	AATTTACAAA GA	1872

ES 2 324 503 T3

(2) INFORMACIÓN DE LA SEQ ID NO: 4:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 5 (A) LONGITUD: 820 pares de bases
 (B) TIPO: ácido nucleico
 (C) NÚMERO DE HEBRAS: doble
 10 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: cDNA

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 4:

15 **AGAAGCTGGC ATCAGAAAAA CAGAGGGGAG ATTTGTGTGG CTGCAGCCGA GGGAGACCAG 60**
GAAGATCTGC ATGGTGGGAA GGACCTGATG ATACAGAGGT GAGAAATAAG AAAGGCTGCT 120
 20 **GACTTTACCA TCTGAGGCCA CACATCTGCT GAAATGGAGA TAATTAACAT CACTAGAAAC 180**
AGCAAGATGA CAATATAATG TCTAAGTAGT GACATGTTTT TGCACATTTC CAGCCCCTTT 240
AAATATCCAC ACACACAGGA AGCACAAAAG GAAGCACAGA GATCCCTGGG AGAAATGCCC 300
 25 **GGCCGCCATC TTGGGTCATC GATGAGCCTC GCCCTGTGCC TGGTCCCGCT TGTGAGGGAA 360**
GGACATTAGA AAATGAATTG ATGTGTTCTT TAAAGGATGG GCAGGAAAAC AGATCCTGTT 420
 30 **GTGGATATTT ATTTGAACGG GATTACAGAT TTGAAATGAA GTCACAAAGT GAGCATTACC 480**
AATGAGAGGA AACAGACGA GAAAATCTTG ATGGCTTCAC AAGACATGCA ACAAACAAAA 540
 35 **TGGAATACTG TGATGACATG AGGCAGCCAA GCTGGGGAGG AGATAACCAC GGGGCAGAGG 600**
GTCAGGATTC TGGCCCTGCT GCCTAACTG TGCCTTCATA ACCAAATCAT TTCATATTTT 660
TAACCCTCAA AACAAAGCTG TTGTAATATC TGATCTCTAC GGTTCTTCT GGGCCCAACA 720
 40 **TTCTCCATAT ATCCAGCCAC ACTCATTTTT AATATTTAGT TCCAGATCT GTACTGTGAC 780**
CTTTCTACAC TGTAGAATAA CATTACTCAT TTTGTTCAAA 820

(2) INFORMACIÓN DE LA SEQ ID NO: 5:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 50 (A) LONGITUD: 11 aminoácidos
 (B) TIPO: aminoácido
 (C) NÚMERO DE HEBRAS: desconocido
 55 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 5:

60 **His Thr Gln Glu Ala Gln Lys Glu Ala Gln Arg**
1 5 10

(2) INFORMACIÓN DE LA SEQ ID NO: 6:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 3582 pares de bases

ES 2 324 503 T3

- (B) TIPO: ácido nucleico
(C) NÚMERO DE HEBRAS: doble
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: cDNA

(ix) CARACTERÍSTICAS:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 401..553

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 6:

15	ACAGAAGAAA TAGCAAGTGC CGAGAAGCTG GCATCAGAAA AACAGAGGGG AGATTTGTGT	60
	GGCTGCAGCC GAGGGAGACC AGGAAGATCT GCATGGTGGG AAGGACCTGA TGATACAGAG	120
20	GAATTACAAC ACATATACTT AGTGTTCCTCA TGAACACCAA GATAAATAAG TGAAGAGCTA	180
	GTCCGCTGTG AGTCTCCTCA GTGACACAGG GCTGGATCAC CATCGACGGC ACTTCTGAG	240
	TACTCAGTGC AGCAAAGAAA GACTACAGAC ATCTCAATGG CAGGGGTGAG AAATAAGAAA	300
25	GGCTGCTGAC TTTACCATCT GAGGCCACAC ATCTGCTGAA ATGGAGATAA TTAACATCAC	360
	TAGAAACAGC AAGATGACAA TATAATGTCT AAGTAGTGAC ATG TTT TTG CAC ATT	415
	Met Phe Leu His Ile	
	55	
30	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	
	60 65 70	
35	CAG AGA TCC CTG GGA GAA ATG CCC GGC CGC CAT CTT GGG TCA TCG ATG	511
	Gln Arg Ser Leu Gly Glu Met Pro Gly Arg His Leu Gly Ser Ser Met	
	75 80 85	
40	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	
	90 95 100	
	TAGAAAATGA ATTGATGTGT TCCTTAAAGG ATGGGCAGGA AACAGATCC TGTTGTGGAT	613
45	ATTTATTGTA ACGGGATTAC AGATTTGAAA TGAAGTCACA AAGTGAGCAT TACCAATGAG	673
	AGGAAAACAG ACGAGAAAAT CTTGATGGCT TCACAAGACA TGCAACAAAC AAAATGGAAT	733
50	ACTGTGATGA CATGAGGCAG CCAAGCTGGG GAGGAGATAA CCACGGGGCA GAGGGTCAGG	793
	ATTCTGGCCC TGCTGCCTAA ACTGTGCGTT CATAACCAA TCATTTTCATA TTTCTAACCC	853
	TCAAAACAAA GCTGTTGTAA TATCTGATCT CTACGGTTCC TTCTGGGCCC AACATTCTCC	913
55	ATATATCCAG CCACACTCAT TTTTAATATT TAGTTCCCAG ATCTGTACTG TGACCTTTCT	973
	ACACTGTAGA ATAACATTAC TCATTTTGTT CAAAGACCCT TCGTGTTGCT GCCTAATATG	1033
60	TAGCTGACTG TTTTTCCTAA GGAGTGTTCT GGCCAGGGG ATCTGTGAAC AGGCTGGGAA	1093
	GCATCTCAAG ATCTTTCCAG GGTATACTT ACTAGCACAC AGCATGATCA TTACGGAGTG	1153

ES 2 324 503 T3

	AATTATCTAA TCAACATCAT CCTCAGTGTC TTTGCCCAT CTGAAATTCA TTTCCCACTT	1213
	TTGTGCCCAT TCTCAAGACC TCAAATGTC ATTCCATTAA TATCACAGGA TTAACTTTTT	1273
5	TTTTTAACCT GGAAGAATTC AATGTTACAT GCAGCTATGG GAATTTAATT ACATATTTTG	1333
	TTTTCCAGTG CAAAGATGAC TAAGTCCTTT ATCCCTCCCC TTGTTTGAT TTTTTTCCA	1393
10	GTATAAAGTT AAAATGCTTA GCCTTGTA CTGAGCTGTAT ACAGCACAGC CTCTCCCCAT	1453
	CCCTCCAGCC TTATCTGTCA TCACCATCAA CCCCTCCCAT ACCACCTAAA CAAATCTAA	1513
	CTTGTAATTC CTTGAACATG TCAGGACATA CATTATTCCT TCTGCCTGAG AAGCTCTTCC	1573
15	TTGTCTCTTA AATCTAGAAT GATGTAAAGT TTTGAATAAG TTGACTATCT TACTTCATGC	1633
	AAAGAAGGGA CACATATGAG ATTCATCATC ACATGAGACA GCAAATACTA AAAGTGTAAT	1693
20	TTGATTATAA GAGTTTAGAT AAATATATGA AATGCAAGAG CCACAGAGGG AATGTTTATG	1753
	GGGCACGTTT GTAAGCCTGG GATGTGAAGC AAAGGCAGGG AACCTCATAG TATCTTATAT	1813
	AATATACTTC ATTTCTCTAT CTCTATCACA ATATCCAACA AGCTTTTCAC AGAATTCATG	1873
25	CAGTGCAAAT CCCCAAAGGT AACCTTTATC CATTTTCATGG TGAGTGCCTT TTAGAATTTT	1933
	GGCAAATCAT ACTGGTCACT TATCTCAACT TTGAGATGTG TTGTCCTTG TAGTTAATTG	1993
30	AAAGAAATAG GGCACCTCTG TGAGCCACTT TAGGGTTCAC TCCTGGCAAT AAAGAATTTA	2053
	CAAAGAGCTA CTCAGGACCA GTTGTTAAGA GCTCTGTGTG TGTGTGTGTG TGTGTGTGAG	2113
	TGTACATGCC AAAGTGTGCC TCTCTCTCTT GACCCATTAT TTCAGACTTA AAACAAGCAT	2173
35	GTTTTCAAAT GGCACATATGA GCTGCCAATG ATGTATCACC ACCATATCTC ATTATTCTCC	2233
	AGTAAATGTG ATAATAATGT CATCTGTAA CATAAAAAA GTTTGACTTC ACAAAGCAG	2293
40	CTGGAAATGG ACAACCACAA TATGCATAAA TCTAACTCCT ACCATCAGCT ACACACTGCT	2353
	TGACATATAT TGTTAGAAGC ACCTCGCATT TGTGGGTTCT CTTAAGCAA ATACTTGAT	2413
	TAGGTCTCAG CTGGGGCTGT GCATCAGGCG GTTTGAGAAA TATTCAATTC TCAGCAGAAG	2473
45	CCAGAATTTG AATTCCTCA TCTTTTAGGA ATCATTACC AGGTTTGGAG AGGATTGAGA	2533
	CAGCTCAGGT GCTTTCATA ATGTCTCTGA ACTTCTGTCC CTCTTGTGT TCATGGATAG	2593
50	TCCAATAAAT AATGTTATCT TTGAAGTAT GCTCATAGGA GAGAATATAA GAAGTCTGAG	2653

ES 2 324 503 T3

```

TGATATCAAC ATTAGGGATT CAAAGAAATA TTAGATTTAA GCTCACACTG GTCAAAAGGA 2713
ACCAAGATAC AAAGAACTCT GAGCTGTCAT CGTCCCCATC TCTGTGAGCC ACAACCAACA 2773
5 GCAGGACCCA ACGCATGTCT GAGATCCTTA AATCAAGGAA ACCAGTGTCA TGAGTTGAAT 2833
TCTCCTATTA TGGATGCTAG CTTCTGGCCA TCTCTGGCTC TCCTCTTGAC ACATATTAGC 2893
10 TTCTAGCCTT TGCTTCCACG ACTTTTATCT TTTCTCCAAC ACATCGCTTA CCAATCCTCT 2953
CTCTGCTCTG TTGCTTTGGA CTTCCCCACA AGAATTTCAA CGACTCTCAA GTCTTTTCTT 3013
CCATCCCCAC CACTAACCTG AATTGCCTAG ACCCTTATTT TTATTAATTT CCAATAGATG 3073
15 CTGCCTATGG GCTAATATTG CTTTAGATGA ACATTAGATA TTTAAAGTCT AAGAGGTTCA 3133
AAATCCAACCT CATTATCTTC TCTTTCTTTC ACCTCCCCTG CTCCTCTCCC TATATTACTG 3193
20 ATTGACTGAA CAGGATGGTC CCCAAGATGC CAGTCAAATG AGAAACCCAG TGGCTCCTTG 3253
TGGATCATGC ATGCAAGACT GCTGAAGCCA GAGGATGACT GATTACGCCT CATGGGTGGA 3313
GGGGACCACT CCTGGGCCTT CGTGATTGTC AGGAGCAAGA CCTGAGATGC TCCCTGCCTT 3373
25 CAGTGTCTCT TGCATCTCCC CTTTCTAATG AAGATCCATA GAATTTGCTA CATTGAGAA 3433
TTCCAATTAG GAACTCACAT GTTTTATCTG CCCTATCAAT TTTTAAACT TGCTGAAAAT 3493
30 TAAGTTTTTT CAAAATCTGT CCTTGTAAT TACTTTTTCT TACAGTGTCT TGGCATACTA 3553
TATCAACTTT GATTCTTTGT TACAACCTT 3582

```

(2) INFORMACIÓN DE LA SEQ ID NO: 7:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 51 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 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

```

(2) INFORMACIÓN DE LA SEQ ID NO: 8:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 74 pares de bases

ES 2 324 503 T3

(B) TIPO: ácido nucleico

(C) NÚMERO DE HEBRAS: doble

(D) TOPOLOGÍA: lineal

5

(ii) TIPO DE MOLÉCULA: DNA (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 8:

10

GGATCCACAG AAGAAATAGC AAGTGCCGAG AAGCTGGCAT CAGAAAAACA GAGGGGAGAT 60

TTGTGTGGCT GCAG 74

15

20

25

30

35

40

45

50

55

60

65