



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 266 245**

(51) Int. Cl.:

C12N 15/57 (2006.01)
C12N 9/64 (2006.01)
C12N 1/21 (2006.01)
C12N 5/10 (2006.01)
C07K 16/40 (2006.01)
C12Q 1/68 (2006.01)
C12Q 1/37 (2006.01)
A61K 48/00 (2006.01)
A61K 39/395 (2006.01)
A61K 31/7088 (2006.01)
A61K 38/48 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **01966837 .5**

(86) Fecha de presentación : **10.08.2001**

(87) Número de publicación de la solicitud: **1309703**

(87) Fecha de publicación de la solicitud: **14.05.2003**

(54) Título: **Gen de la calicreína.**

(30) Prioridad: **11.08.2000 US 224853 P**

(45) Fecha de publicación de la mención BOPI:
01.03.2007

(45) Fecha de la publicación del folleto de la patente:
01.03.2007

(73) Titular/es: **MOUNT SINAI HOSPITAL**
600 University Avenue
Toronto, Ontario M5G 1X5, CA

(72) Inventor/es: **Yousef, George M. y**
Diamandis, Eleftherios P.

(74) Agente: **Curell Suñol, Marcelino**

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

Gen de la calicreína.

5 **Campo de la invención**

La invención se refiere a moléculas de ácido nucleico, a proteínas codificadas por dichas moléculas de ácido nucleico; y a la utilización de las proteínas y de las moléculas de ácido nucleico.

10 **Antecedentes de la invención**

Las calicreínas son un grupo de serina proteasas que se encuentran en diversos tejidos y fluidos biológicos. El término “calicreína” fue introducido por primera vez por Werle y colegas que descubrió concentraciones elevadas de sus cepas originales en el páncreas (en griego, “kallikreas”) (1, 2). Las calicreínas se dividen en dos grupos principales; la calicreína del plasma, que es un gen individual (3), y las calicreínas tisulares, que están codificadas por una gran familia de multigenes en roedores (4, 5). Hasta hace poco, la familia del gen de la calicreína humana se pensaba que estaba constituida únicamente por tres miembros (6). Sin embargo, se han identificado 11 nuevos miembros de la familia del gen de la calicreína (7 a 18). La evolución en esta área de investigación se ha estudiado recientemente (7).

El antígeno específico de la próstata (PSA), actualmente el marcador tumoral más útil para el diagnóstico y control del cáncer de próstata, es un miembro de la familia del gen de la calicreína humana de las serinas proteasas (19, 20). Además del PSA, la calicreína 2 glandular humana (hK2, codificada por el gen KLK2) se ha propuesto como un marcador adyuvante del diagnóstico para el cáncer de próstata (21, 22). Además, las pruebas acumuladas indican que otros miembros de la familia del gen de la calicreína ampliada pueden estar asociados con el cáncer (7). El gen 1 específico de las células epiteliales normal (NES1) (KLK10, según la nomenclatura del gen de la calicreína del tejido humano aprobado) se descubrió que era un nuevo supresor tumoral, que está regulado por disminución durante la evolución del cáncer de mama (23). Se descubrió asimismo que otros elementos de la familia del gen, incluyendo zima (KLK6), neuropsina (KLK8) y la enzima quimotriptica de la capa córnea humana (HSCCE; KLK7) se expresaban de forma diferenciada en determinados tipos de cánceres (24-26).

30 **Sumario de la invención**

Los presentes inventores identificaron una molécula de ácido nucleico que codifica una nueva calicreína. La molécula de ácido nucleico cartografía al cromosoma 19q13.3-q13.4 y está situado entre los genes *klk1* y *klk3*. La nueva molécula de ácido nucleico denominada “*klk15*” presenta tres formas de corte y empalme alternativamente y se expresa principalmente en la glándula tiroides, y en menor extensión en la próstata, glándulas salivales y adrenales, colon, testículos y riñón. La expresión del ácido nucleico está regulada por incremento en el cáncer de próstata y está bajo regulación de la hormona esteroide en la línea celular LNCaP del cáncer de próstata. La elevada expresión de *klk15* está asociada a los tumores de próstata más agresivos (etapa superior y grado superior).

La nueva proteína calicreína descrita en la presente memoria se denomina “Calicreína 15”, “KLK15” o “proteína de KLK15”. El gen que codifica la proteína se denomina “*klk15*”.

La presente invención se refiere a una proteína aislada que está constituida por una secuencia de aminoácidos de la SEC. ID n°: 6, 7, 8 ó 9.

La presente invención se refiere además a una molécula de ácido nucleico aislada que codifica dicha proteína o a una secuencia de ácidos nucleicos complementaria a ésta o a una molécula de ácido nucleico que presenta por lo menos 80%, 85%, 90%, 95% o 99% de identidad con ésta.

La presente invención se refiere además a una molécula aislada de ácido nucleico que presenta una secuencia de ácidos nucleicos de la SEC. ID n°: 1, 2, 3, 4 ó 5 en la que T puede además ser U, o una secuencia de ácidos nucleicos complementaria a ésta, o una molécula de ácido nucleico que presenta por lo menos 80%, 85%, 90%, 95% o 99% de identidad con ésta.

Las moléculas de ácido nucleico de la invención pueden insertarse en un vector de expresión apropiado, es decir, un vector que contiene los elementos necesarios para la transcripción y traducción de la secuencia de codificación insertada. Por consiguiente, pueden construirse vectores de expresión recombinante adaptados para la transformación de una célula huésped que comprenden una molécula de ácido nucleico de la invención y uno o más elementos de transcripción y traducción unidos a la molécula de ácido nucleico.

El vector de expresión recombinante puede utilizarse para preparar células huésped transformadas que expresan proteínas de KLK15. Por consiguiente, la invención proporciona además células huésped que contienen una molécula recombinante de la invención. La invención contempla asimismo mamíferos no humanos transgénicos cuyas células germinales y células somáticas contienen una molécula recombinante que comprende una molécula de ácido nucleico de la invención, en particular la que codifica un análogo de la proteína de KLK15 o un truncamiento de la proteína de KLK15.

La invención proporciona además un método para preparar proteínas de KLK15 que utiliza las moléculas de ácido nucleico purificadas y aisladas de la invención. En una forma de realización se proporciona un método para la preparación de una proteína de KLK15 que comprende (a) transferir un vector de expresión recombinante de la invención a una célula huésped; (b) seleccionar las células huésped transformadas de entre las células huésped no transformadas; (c) cultivar una célula huésped transformada seleccionada en condiciones que permitan la expresión de la proteína de KLK15; y (d) aislar la proteína de KLK15.

Las proteínas de KLK15 de la invención pueden conjugarse con otras moléculas, tales como las proteínas, para preparar proteínas de fusión. Esto puede realizarse, por ejemplo, mediante la síntesis de proteínas de fusión N-terminales o C-terminales.

La invención contempla además anticuerpos que se unen específicamente a una proteína de KLK15 de la invención. Los anticuerpos pueden marcarse con una sustancia detectable y utilizarse para detectar las proteínas de la invención en tejidos y células. Los anticuerpos pueden presentar utilización particular en aplicaciones terapéuticas, por ejemplo para reaccionar con células tumorales y en conjugados e inmunotoxinas como portadores selectivos de la diana de varios agentes que presentan efectos antitumorales incluyendo fármacos quimioterapéuticos, toxinas, modificadores de la respuesta inmunológicas, enzimas y radioisótopos.

La invención permite además la construcción de sondas de nucleótido que son exclusivas para las moléculas de ácido nucleico de la invención y/o para las proteínas de la invención. Por consiguiente, la invención se refiere también a una sonda que comprende una secuencia de ácidos nucleicos de la invención o a una secuencia de ácidos nucleicos que codifican una proteína de la invención o una parte de ésta. La sonda puede marcarse, por ejemplo, con una sustancia detectable y puede utilizarse para seleccionar de entre una mezcla de secuencias de nucleótidos una molécula de ácido nucleico de la invención incluyendo las moléculas de ácido nucleico que codifican una proteína que presenta una o más propiedades de una proteína de la invención. Puede utilizarse una sonda para marcar los tumores.

La invención proporciona también moléculas de ácido nucleico con cadena complementaria, p. ej. para la producción de una cadena de ARNm o de ADN en orientación inversa a una molécula de cadena complementaria. Una molécula de ácido nucleico con cadena complementaria puede utilizarse para suprimir el crecimiento de una célula (p. ej. cancerosa) que expresa KLK15. La invención proporciona además todavía un método para identificar una sustancia que se une a una proteína de la invención que comprende hacer reaccionar la proteína con al menos una sustancia que se une potencialmente con la proteína, en condiciones que permiten la formación de complejos entre la sustancia y la proteína y que detecta el enlace. El enlace puede detectarse analizando los complejos, la sustancia libre o la proteína no acomplexada. La invención contempla también los métodos de identificación de sustancias que se unen a otras proteínas intracelulares que interactúan con una proteína de KLK15. Pueden utilizarse también métodos para identificar los compuestos que se unen a secuencias reguladoras del gen KLK15 (p. ej. secuencias activadoras).

Más aún la invención proporciona un método para evaluar la capacidad de un compuesto para modular la actividad biológica de una proteína de KLK15 de la invención. Por ejemplo puede evaluarse una sustancia que inhibe o aumenta la interacción de la proteína y una sustancia que se une a la proteína. En una forma de realización, el método comprende proporcionar una concentración conocida de una proteína de KLK15, con una sustancia que se une a la proteína y un compuesto de análisis en condiciones que permitan la formación de complejos entre la sustancia y la proteína y la eliminación y/o detección de complejos.

Los compuestos que modulan la actividad biológica de una proteína de la invención pueden identificarse también utilizando los métodos de la invención comparando el patrón y el nivel de expresión de la proteína de la invención en tejidos y células, en presencia y ausencia de los compuestos.

Las proteínas de la invención, los anticuerpos, las moléculas de ácido nucleico con cadena complementaria y las sustancias y compuestos identificados utilizando los métodos de la invención, y los péptidos de la invención pueden utilizarse para modular la actividad biológica de una proteína de KLK15 de la invención, y pueden utilizarse en el tratamiento de enfermedades tales como el cáncer (particularmente el cáncer de próstata, de colon, de riñón y testicular) y trastornos de tiroides en un paciente. Por consiguiente, las sustancias y compuestos pueden formularse en composiciones para la administración a individuos que padecen trastornos tales como el cáncer (particularmente el cáncer de próstata, colon, riñón y testicular) y trastornos del tiroides en un paciente. En particular, pueden utilizarse los anticuerpos, las moléculas de ácido nucleico con cadena complementaria, las sustancias y los compuestos para tratar pacientes que tienen una proteína de KLK15 en sus células cancerosas o sobre ellas.

Por consiguiente, la presente invención se refiere también a una composición que comprende una o más moléculas de ácido nucleico o a una proteína de la invención, y un portador, excipiente o diluyente farmacéuticamente aceptable. Se proporciona también la utilización de una molécula de ácido nucleico o proteína de la invención para la preparación de un medicamento destinado a su utilización en un método de tratamiento o prevención de un trastorno tal como el cáncer (particularmente el cáncer de próstata, de tiroides, de colon, de riñón y testicular) y trastornos del tiroides en un paciente.

Otro aspecto de la invención consiste en la utilización de una proteína de KLK15 o de su molécula de ácido nucleico codificadora en la preparación de vacunas para prevenir el cáncer y/o tratar el cáncer, en particular para prevenir y/o

tratar el cáncer en pacientes que tienen una proteína de KLK15 detectada en sus células. Estas preparaciones de vacuna pueden utilizarse también para impedir que los pacientes presenten tumores antes de su aparición.

La invención contempla ampliamente las vacunas para estimular o potenciar la producción de anticuerpos dirigidos contra una proteína de KLK15 en un paciente al que se administra la vacuna.

La invención proporciona además la utilización de una proteína de la invención para la preparación de una vacuna para estimular o potenciar en un paciente la producción de anticuerpos dirigida contra una proteína de KLK15. El método comprende la administración a un paciente de una vacuna de la invención en una dosis eficaz para estimular o aumentar la producción de anticuerpos.

La vacuna puede utilizarse en métodos para el tratamiento, prevención o retardo de la aparición del cáncer. Los métodos comprenden administrar a un sujeto una vacuna de la invención en una dosis eficaz para tratar, prevenir o retardar la aparición del cáncer.

En otras formas de realización, la invención proporciona un método para identificar inhibidores de una interacción de la proteína de KLK15, que comprende

- (a) proporcionar una mezcla de reacción que incluye la proteína de KLK15 y una sustancia que se une a la proteína de KLK15, o por lo menos una parte de cada una que interactúa;
- (b) poner en contacto la mezcla de reacción con uno o más compuestos de análisis;
- (c) identificar los compuestos que inhiben la interacción de la proteína de KLK15 y la sustancia.

En determinadas formas de realización preferidas, la mezcla de reacción es una célula completa. En otras formas de realización, la mezcla de reacción es un lisado celular o una composición proteica purificada. El método del asunto puede realizarse utilizando bancos de compuestos de ensayo. Dichos agentes pueden ser bancos de proteínas, péptidos, ácidos nucleicos, carbohidratos, pequeñas moléculas orgánicas y extractos de productos naturales, tales como los aislados en animales, plantas, hongos y/o microbios. Otro aspecto aún de la presente invención proporciona un método para realizar el descubrimiento del fármaco que comprende:

- (a) proporcionar uno o más sistemas de ensayo para identificar los agentes por su capacidad para inhibir o potenciar la interacción de una proteína de KLK15 y una sustancia que se une a la proteína;
- (b) realizar el perfil terapéutico de los agentes identificados en la etapa (a) o más análogos de los mismos, la eficacia y toxicidad en animales; y
- (c) determinar una formulación para una preparación farmacéutica que incluye uno o más agentes identificados en la etapa (b) que presentan un perfil terapéutico aceptable.

Otros objetivos, características y ventajas de la presente invención resultarán evidentes a partir de la siguiente descripción detallada. Debe entenderse, sin embargo, que la descripción detallada y los ejemplos específicos aunque indican formas de realización preferidas de la invención se proporcionan únicamente a título ilustrativo.

Breve descripción de los dibujos

La invención se describe a continuación haciendo referencia a los dibujos en los que:

La Figura 1 presenta la organización genómica y la secuencia genómica parcial del gen KLK15. Las secuencias intrónicas no se presentan excepto para las áreas de unión de corte y empalme. Se presentan los intrones con letras minúsculas y los exones con letras mayúsculas. Los nucleótidos de codificación se presentan en negrita y la zona 3' no traducida sigue al codón de terminación TGA (rodeado por un círculo). Los aminoácidos traducidos de la zona de codificación se presentan bajo una sola letra de abreviatura. Los codones de iniciación y terminación están rodeados por un círculo y las uniones exón-intrón están subrayadas. Los restos catalíticos están encasillados. La supuesta señal de poliadenilación está subrayada. El punto de partida exacto del primer exón de codificación no se determinó.

La Figura 2 presenta una alineación de la secuencia deducida de aminoácidos de KLK15 con los miembros de la familia del multigén con caliceína (SEC. ID n°: 25 a 38). Los guiones representan huecos a los que llevar las secuencias para mejor alineación. Los restos de la triada catalítica (H, D, S) se presentan en cursiva. Los aminoácidos idénticos están destacados en negrita y las secuencias similares en gris. Los 29 restos de serina proteasa invariable están marcados por (●) en la parte inferior y los restos de cisteína por (+) en la parte superior de cada bloque. Las secuencias de escisión previstas de los péptidos señal y de activación están indicadas con flechas. El área punteada representa la secuencia del bucle de caliceína. El modelo de escisión de tipo tripsina previsto por la presencia del resto "D" está indicado por (*). KLK15 presenta una "E" en esta posición. Una única secuencia de 8 aminoácidos, HNEPGTAG (SEC. ID n°: 10), está presente en las posiciones 148 a 155 del gen KLK15.

La Figura 3 es una representación de la hidrofobia y hidrofilia de la proteína de KLK15, en comparación con el antígeno específico de la próstata (PSA). Obsérvese la zona hidrófoba en el terminal amino, que sugiere la presencia de un péptido señal.

La Figura 4 es un dendograma del árbol filogenético previsto para 15 calicreínas y otras pocas serina proteasas. Se utilizó el método de unión próxima para alinear KLK15 con otras serinas proteasas y miembros de la familia del gen de la calicreína. El árbol agrupó las calicreínas clásicas (hK1, hK2 y PSA) y alineó KLK15 en un grupo con los genes TLSP y KLK-L3. Otras serina proteasas se alinearon en diferentes grupos, como se muestra. KLK representa calicreína; KLK-L representa pseudo calicreína; TLSP representa pseudotripsina serina proteasa; NES1 representa el gen normal específico de las células epiteliales; PSA representa el antígeno específico de la próstata; hK1 y hK2 representan la calicreína 1 y 2 glandular humana, respectivamente; y HSCCE representa la enzima quimotriptica de la capa córnea humana.

La Figura 5 es una presentación esquemática de las variantes de corte y empalme diferente de un gen con KLK15. Los exones se presentan encasillados y los intrones mediante líneas de conexión. Los números dentro de las casillas representan las longitudes del exón en pares de bases. La punta de flecha apunta al codón de iniciación común y comienza en las posiciones del codón de terminación. La longitud del producto del polipéptido previsto está indicada junto a cada variante en los aminoácidos (AA). El corte y empalme alternativo y/o las omisiones del exón crean un desplazamiento del marco, que conduce a una terminación prematura.

La Figura 6 presenta las posiciones relativas de los genes con KLK1, KLK15 y KLK3 en el cromosoma 19q13.3-q13.4. Los dos clones BAC solapantes se identifican y la zona de solapamiento está sombreada. Los genes están representados mediante flechas horizontales que indican la dirección de la secuencia de codificación. Las distancias entre los genes se mencionan en pares de bases. La Figura no está dibujada a escala.

La Figura 7 presenta la expresión tisular del gen con KLK15, determinada por RT-PCR. KLK15 se expresa principalmente en la glándula tiroides y en menor extensión en la próstata, glándulas salivales y adrenales, en el colon, testículos y riñón. M = marcador de peso molecular. Para la explicación de las bandas múltiples de PCR (como alternativas formas de corte y empalme) véase el Ejemplo. Se realizó la PCR con los cebadores KLK15-F2 y KLK15-R1.

La Figura 8 presenta la regulación hormonal del gen con KLK15 en la línea celular LNCaP de cáncer de próstata. DHT = dihidrotestosterona. Se añadieron esteroides a una concentración final de 10^{-8} M. (-ve) = referencia negativa. Se utilizó actina como gen de referencia.

La Figura 9 es un diagrama esquemático que presenta la comparación de las zonas de codificación de los genes de la calicreína 15. Los exones están representados mediante barras negras y los intrones mediante las líneas de conexión. Las letras por encima de las casillas indican las posiciones relativas de la triada catalítica que se descubrió que se conservaba en todos los genes; H indica histidina, D ácido aspártico y S serina. Los números romanos indican las fases de intrón. La fase de intrón se refiere a la posición del intrón dentro del codón; I indica que el intrón aparece después del primer nucleótido del codón, II el intrón aparece después del segundo nucleótido, 0 que el intrón aparece entre los codones. Las fases de intrón se conservan en todos los genes. Los números dentro de las casillas indican longitudes del exón en pares de bases. Los nombres entre paréntesis representan la nomenclatura oficial aprobada por el comité de nomenclatura génica humana. Las zonas 3' y 5' no traducidas y los exones en 5' no traducidos no se presentan.

Descripción detallada de la invención

Según la presente invención puede emplearse la biología molecular convencional, la microbiología y las técnicas de ADN recombinante dentro de los conocimientos en la materia. Dichas técnicas se explican completamente en la bibliografía. Véase por ejemplo, Sambrook, Fritsch y Maniatis, *Molecular Cloning: A Laboratory Manual*, segunda edición (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; *DNA Cloning: A Practical Approach*, volúmenes I y II (D.N. Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed. 1984); *Nucleic Acid Hybridization* B.D. Hames y S. J. Higgins eds. (1985); *Transcription and Translation* B. D. Hames y S.I. Higgins eds (1984); *Animal Cell Culture* R. I. Freshney, ed. (1986); *Immobilized Cells and enzymes* IRL Press, (1986); y B. Perbal, *A Practical Guide to Molecular Cloning* (1984).

1. Moléculas de ácido nucleico de la invención

Tal como se mencionó en la presente memoria anteriormente, la invención proporciona una molécula aislada de ácido nucleico que presenta la secuencia que codifica una proteína de KLK15. El término "aislado" se refiere a un ácido nucleico sustancialmente exento de material celular o al medio de cultivo cuando se produce por técnicas de ADN recombinante o a reactivos químicos o a otros productos químicos cuando se sintetizan químicamente. Un ácido nucleico "aislado" puede también estar exento de las secuencias que flanquean de forma natural al ácido nucleico (es decir, las secuencias situadas en los extremos 5' y 3' de la molécula de ácido nucleico) de la que procede el ácido nucleico. La expresión "ácido nucleico" pretende incluir el ADN y ARN y puede ser de doble cadena o de una cadena. Las moléculas de ácido nucleico de la invención codifican una proteína que está constituida por una secuencia de aminoácidos de la SEC. ID nº: 6, 7, 8 ó 9. Preferentemente las moléculas de ácido nucleico de la invención comprenden una secuencia de ácido nucleico de una o más de las SEC. ID nº: 1 a 5.

La invención incluye secuencias de ácido nucleico complementarias a un ácido nucleico que codifican una proteína que está constituida por una secuencia de aminoácidos de la SEC. ID nº: 6, 7, 8 ó 9 preferentemente las secuencias de ácidos nucleicos complementarias a una secuencia de ácido nucleico completa de una o más de las SEC. ID nº: 1 a 5.

5 La invención incluye moléculas de ácido nucleico con identidad y homología de la secuencia sustancial de las secuencias de ácido nucleico de la invención o proteínas codificantes con identidad o similitud sustancial a la secuencia de aminoácidos de las SEC. ID nº: 6, 7, 8 ó 9. Preferentemente, los ácidos nucleicos presentan identidad sustancial de secuencia por ejemplo como mínimo del 80% o 85% de identidad del ácido nucleico; más preferentemente del 90% de identidad de ácido nucleico; y aún más preferentemente por lo menos del 95%, 96%, 97%, 98% ó 99% de identidad de secuencia. "Identidad" como se conoce en la técnica y se utiliza en la presente memoria, es una
10 relación entre dos o más secuencias de aminoácidos o dos o más secuencias de ácidos nucleicos, determinadas por comparación de las secuencias. Se refiere también al grado de relevancia entre las secuencias de aminoácido o ácido nucleico, como puede ser el caso, determinado por la compatibilidad entre las cadenas de cada secuencia. Identidad y similitud son términos bien conocidos por los expertos en la materia y pueden calcularse por métodos convencionales (por ejemplo véase Computational Molecular Biology, Lesk, A. M. ed., Oxford University Press, Nueva York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W. ed., Academic Press, Nueva York, 1993; Computer
15 Analysis of Sequence Data, Part I, Griffin, A. M. y Griffin, H. G. eds., Humana Press, Nueva Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G. Academic Press, 1987; y Sequence Analysis Primer, Gribskov, M. y Devereux, J. eds. M. Stockton Press, Nueva York, 1991, Carillo, H. y Lipman, D., SIAM J. Applied Math. 48: 1073, 1988). Se prefieren generalmente los métodos que están diseñados para proporcionar la mayor compatibilidad entre las dos secuencias. Los métodos para determinar la identidad y la similitud están codificados en programas informáticos disponibles al público que incluyen el paquete del programa CGC (Devereux J. *et al.*, *Nucleic Acids Research* 12(1): 387, 1984); BLASTP, BLASTN y FASTA (Altschul, S. F. *et al.*, *J. Molec. Biol.* 215: 403-410, 1990). El programa BLAST X está disponible al público en NCBI y otras fuentes (BLAST Manual, Altschul, S. *et al.* NCBI NLM NIH
20 Bethesda, Md. 20894; Altschul, S. *et al.* *J. Mol. Biol.* 215: 403-410, 1990).

Las moléculas de ácido nucleico aisladas que codifican una proteína de KLK15 y las que presentan una secuencia que se diferencia de una secuencia de ácido nucleico de la invención debido a la degeneración del código genético están también comprendidas en el alcance de la invención. Dichos ácidos nucleicos codifican proteínas funcionalmente
30 equivalentes (p. ej. una proteína de KLK15) pero se diferencian en la secuencia procedente de la secuencia de una proteína de KLK15 debido a la degeneración del código genético. Como ejemplo, los polimorfismos de la secuencia de ADN en la secuencia de nucleótidos de una proteína de KLK15 pueden producir mutaciones imperceptibles que no afectan a la secuencia de aminoácidos. Pueden existir variaciones en uno o más nucleótidos entre los individuos en una población debido a la variación alélica natural. Cualquiera y todas las variaciones de ácidos nucleicos están
35 comprendidas dentro del alcance de la invención. Pueden producirse también polimorfismos de la secuencia de ADN los cuales conducen a cambios en la secuencia de aminoácidos de una proteína de KLK15. Estos polimorfismos de aminoácidos están también comprendidos dentro del alcance de la presente invención.

Una molécula aislada de ácido nucleico de la invención que comprende ADN puede aislarse preparando una sonda
40 marcada de ácido nucleico basada en toda o parte de la secuencia de ácido nucleico de la invención. La sonda de ácido nucleico marcada se utiliza para identificar un banco de ADN apropiado (p. ej. un banco de ADNc o de ADN genómico). Por ejemplo, puede utilizarse un banco de ADNc para aislar un ADNc que codifica una proteína de KLK15 cribando el banco con la sonda marcada utilizando técnicas normalizadas. Como alternativa, puede cribarse asimismo un banco de ADN genómico para aislar un clon genómico que comprende un gen que codifica una proteína de KLK15.
45 Los ácidos nucleicos aislados mediante cribado de un ADNc o de un banco de ADN genómico pueden secuenciarse por técnicas normalizadas. Una molécula de ácido nucleico aislada de la invención que es ADN puede también aislarse ampliando de manera selectiva un ácido nucleico que codifica una proteína de KLK15 utilizando los métodos de la reacción en cadena de polimerasa (PCR) y ADNc o ADN genómico. Es posible diseñar cebadores de oligonucleótido sintéticos a partir de la secuencia de nucleótidos de la invención para su utilización en PCR. Un ácido nucleico puede
50 ampliarse a partir del ADNc o del ADN genómico utilizando estos cebadores de oligonucleótido y técnicas de ampliación por PCR normalizadas. El ácido nucleico ampliado de este modo puede clonarse en un vector apropiado y caracterizarse mediante el análisis de la secuencia del ADN. El ADNc puede prepararse a partir del ARNm, aislando el ARNm celular completo mediante varias técnicas, por ejemplo, utilizando el procedimiento de extracción con tiocianato de guanidinio de Chirgwin *et al.*, *Biochemistry*, 18, 5294-5299 (1979). El ADNc se sintetiza a continuación a partir
55 del ARNm utilizando transcriptasa inversa (por ejemplo, la transcriptasa inversa MLV de Moloney de Gibco/BRL, Bethesda, MD, o transcriptasa inversa AMV disponible en Seikagaku America, Inc., San Petersburgo, FL).

Puede aislarse una molécula aislada de ácido nucleico de la invención que es ARN clonando un ADNc que codifica una proteína de KLK15 en un vector apropiado que permite la transcripción del ADNc para producir una molécula de
60 ARN que codifica una proteína de KLK15. Por ejemplo, un ADNc puede clonarse corriente abajo de un activador de bacteriófago (p. ej. un activador T7) en un vector, el ADNc puede transcribirse *in vitro* con T7 polimerasa y el ARN resultante puede aislarse por técnicas convencionales.

Las moléculas de ácido nucleico de la invención pueden sintetizarse químicamente utilizando las técnicas norma-
65 lizadas. Son conocidos los métodos de síntesis química de polidesoxinucleótidos, que incluyen pero no se limitan a la síntesis en fase sólida que, como la síntesis de péptidos, se ha automatizado completamente en los sintetizadores de ADN disponibles en el mercado (véase p. ej., Itakura *et al.* patente US nº 4.598.049; Caruthers *et al.* patente US nº 4.458.066; e Itakura patentes US nº 4.401.796 y nº 4.373.071).

La determinación de si una molécula determinada de ácido nucleico codifica una proteína de KLK15 puede llevarse a cabo expresando el ADNc en una célula huésped apropiada por técnicas normalizadas y analizando la proteína expresada en los métodos descritos en la presente memoria. Un ADNc que codifica una proteína de KLK15 puede secuenciarse por técnicas normalizadas, tal como la terminación de la cadena de didesoxinucleótido o el secuenciado químico de Maxam-Gilbert, para determinar la secuencia de ácidos nucleicos y la secuencia de aminoácidos prevista de la proteína codificada.

El codón de iniciación y las secuencias no traducidas de una proteína de KLK15 pueden determinarse utilizando el programa informático diseñado para este fin, tal como PC/Gene (IntelliGenetics Inc., Calif.). La estructura intrón-exón y las secuencias reguladoras de la transcripción de un gen que codifica una proteína de KLK15 pueden confirmarse utilizando una molécula de ácido nucleico de la invención que codifica una proteína de KLK15 para sondear un banco de clones de ADN genómico. Pueden identificarse elementos reguladores utilizando técnicas normalizadas. La función de los elementos puede confirmarse utilizando estos elementos para expresar un gen indicador tal como el gen lacZ que está operativamente unido a los elementos. Estos montajes pueden introducirse en células cultivadas utilizando procedimientos convencionales o en modelos de animales transgénicos no humanos. Además de identificar elementos reguladores en el ADN, dichos montajes pueden utilizarse también para identificar proteínas nucleares que interactúan con los elementos, utilizando técnicas conocidas en la materia.

En una determinada forma de realización de la invención, las moléculas de ácido nucleico aisladas utilizando los métodos descritos en la presente memoria son alelos del gen con KLK15 mutante. Los alelos mutantes pueden aislarse de individuos ya sean conocidos o propuestos que presenten un genotipo que contribuya a los síntomas de un trastorno que involucra una proteína de KLK15. Los alelos mutantes y los productos de alelo mutante pueden utilizarse en métodos terapéuticos y de diagnóstico descritos en la presente memoria. Por ejemplo, un ADNc de un gen mutante con KLK15 puede aislarse utilizando PCR tal como se describe en la presente memoria, y la secuencia del ADN del alelo mutante puede compararse con la del alelo normal para determinar la(s) mutación(es) responsable(s) de la pérdida o alteración de la función del producto génico mutante. Asimismo puede construirse un banco genómico utilizando ADN de un individuo que se sospecha o se conoce que lleva un alelo mutante, o puede construirse un banco de ADNc utilizando ARN procedente del tejido que se conoce o se sospecha que expresa el alelo mutante. Un ácido nucleico que codifica un gen normal con KLK15 o cualquier fragmento adecuado del mismo, puede marcarse a continuación y utilizarse como sonda para identificar el alelo mutante correspondiente en dichos bancos. Los clones que contienen secuencias mutantes pueden purificarse y someterse a análisis de secuencia. Además, puede construirse un banco de expresión utilizando ADNc procedente del ARN aislado de un tejido de un individuo que se conoce o se sospecha para expresar un alelo mutante de KLK15. Los productos génicos preparados mediante el tejido supuestamente mutante pueden expresarse y cribarse, por ejemplo utilizando anticuerpos específicos para una proteína de KLK15 como se describe en la presente memoria. Los clones del banco identificados que utilizan los anticuerpos pueden purificarse y someterse al análisis de la secuencia.

La secuencia de una molécula de ácido nucleico de la invención, o un fragmento de la molécula, pueden invertirse con relación a su presentación normal para que la transcripción produzca una molécula de ácido nucleico con cadena complementaria. Puede construirse una molécula de ácido nucleico con cadena complementaria utilizando síntesis química o reacciones de ligadura enzimática que utilizan los procedimientos conocidos en la materia.

2. Proteínas de la invención

Una secuencia de aminoácidos de la proteína de KLK15 está constituida por una secuencia como se presenta en las SEC. ID n°: 6, 7, 8 y 9. La proteína se expresa principalmente en la glándula tiroidea y en menor medida en la próstata, glándulas salivales y suprarrenales, colon, testículos y riñón.

La invención contempla asimismo isoformas de las proteínas de la invención. Una isoforma contiene el mismo número y clase de aminoácidos que una proteína de la invención, pero la isoforma presenta una estructura molecular diferente. Las isoformas contempladas por la presente invención presentan preferentemente las mismas propiedades que una proteína de la invención descrita en la presente memoria.

La presente invención incluye también proteínas de KLK15 conjugadas con una proteína, o una proteína marcadora seleccionada (véase a continuación) para producir proteínas de fusión.

Una proteína de KLK15 de la invención puede prepararse utilizando métodos de ADN recombinante. Por consiguiente, las moléculas de ácido nucleico de la presente invención que presentan una secuencia que codifica una proteína de KLK15 de la invención pueden incorporarse de forma conocida en un vector de expresión apropiado que asegura la expresión buena de una proteína. Los vectores de expresión posibles incluyen pero no se limitan a cósmidos, plásmidos o virus modificados (p. ej., retrovirus con defectos de replicación, adenovirus y virus adeno-asociados) siempre que el vector sea compatible con la célula huésped utilizada.

La invención contempla por lo tanto un vector de expresión recombinante de la invención que contiene una molécula de ácido nucleico de la invención y las secuencias reguladoras necesarias para la transcripción y traducción de la secuencia de proteína insertada. Las secuencias reguladoras adecuadas pueden proceder de varias fuentes, incluyendo los genes bacterianos, fúngicos, víricos, de mamífero o de insecto [por ejemplo, véase las secuencias reguladoras descritas en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA

(1990)]. La selección de secuencias reguladoras apropiadas depende de la célula huésped seleccionada como se expone a continuación y un experto en la materia puede realizarla fácilmente. Las secuencias reguladoras necesarias pueden ser suministradas por la proteína de KLK15 natural y sus zonas flanqueantes.

La invención proporciona además un vector de expresión recombinante que comprende una molécula de ácido nucleico ADN de la invención clonada en el vector de expresión en una orientación de cadena complementaria. Esto es, la molécula de ADN está unida a la secuencia reguladora de manera que permite la expresión, mediante transcripción de la molécula de ADN, de una molécula de ARN que es de cadena complementaria con respecto a la secuencia de ácido nucleico de una proteína de la invención o un fragmento de la misma. Pueden seleccionarse secuencias reguladoras unidas al ácido nucleico con cadena complementaria, las cuales dirigen la expresión continua de la molécula de ARN con cadena complementaria en varios tipos de célula, por ejemplo un activador y/o potenciador vírico, o pueden seleccionarse secuencias reguladoras, las cuales dirigen la expresión específica de tipo tisular o celular del ARN con cadena complementaria.

Los vectores con expresión recombinante de la invención pueden también contener un gen marcador que facilita la selección de las células huésped transformadas o transfectadas con una molécula recombinante de la invención. Ejemplos de genes marcadores son los genes que codifican una proteína tal como G418 e higromicina que confieren resistencia a determinados fármacos, β -galactosidasa, cloranfenicol acetiltransferasas, luciferasa de luciérnaga o una inmunoglobulina o fragmento de la misma tal como el fragmento Fc de una inmunoglobulina, preferentemente IgG. Los marcadores pueden introducirse en un vector separado del ácido nucleico de interés.

Los vectores con expresión recombinante pueden también contener genes que codifican un resto de fusión que proporciona aumento de expresión de la proteína recombinante; aumento de solubilidad de la proteína recombinante y ayudan a la purificación de la proteína recombinante diana actuando como ligando en la purificación por afinidad. Por ejemplo, puede añadirse una secuencia de escisión proteolítica a la proteína recombinante diana para que permita la separación de la proteína recombinante del resto de fusión después de la purificación de la proteína de fusión. Los vectores de expresión de la fusión típicos incluyen pGEX (Amrad Corp., Melbourne, Australia), pMAL (New England Biolabs, Beverly, MA) y pRIT5 (Farmacia, Piscataway, NJ) que fusionan la glutatión S-transferasa (GST), proteína de unión a maltosa E o proteína A, respectivamente, a la proteína recombinante.

Pueden introducirse vectores con expresión recombinante en las células huésped para producir una célula huésped transformante. Las "células huésped transformantes" incluyen las células huésped que han sido transformadas o transfectadas con un vector con expresión recombinante de la invención. Las expresiones "transformada con", "transfectada con", "transformación" y "transfección" comprenden la introducción de un ácido nucleico (p. ej. un vector) en una célula mediante una de las muchas técnicas normalizadas. Las células procarióticas pueden transformarse con un ácido nucleico, por ejemplo, por electroporación o transformación mediada por cloruro de calcio. Puede introducirse un ácido nucleico en células de mamífero por técnicas convencionales tales como coprecipitación con fosfato cálcico o cloruro de calcio, transfección mediada por DEAE-dextrano, lipofectina, electroporación o microinyección. Los métodos adecuados para transformar y transfectar células huésped pueden encontrarse en Sambrook *et al.* (*Molecular Cloning: A Laboratory Manual*, 2ª edición, Cold Spring Harbor Laboratory press (1989)), y otros libros de texto de laboratorio.

Las células huésped adecuadas incluyen una amplia variedad de células huésped procarióticas y eucarióticas. Por ejemplo, las proteínas de la invención pueden expresarse en células bacterianas tales como *E. coli*, células de insecto (utilizando baculovirus), células de levadura o células de mamífero. Otras células huésped adecuadas pueden encontrarse en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1991).

Puede seleccionarse también una célula huésped que module la expresión de una secuencia de ácido nucleico insertada o que modifique (p. ej. glucosilación o fosforilación) y procese (p. ej. escinda) la proteína de una manera deseada. Pueden seleccionarse sistemas huésped o líneas celulares para que tengan mecanismos específicos y característicos para el tratamiento después de la traducción y la modificación de las proteínas. Por ejemplo, pueden utilizarse las células huésped eucarióticas que incluyen CHO, VERO, BHK, HeLA, COS, MDCK, 293, 3T3 y WI38. Para la expresión de la proteína estable de alto rendimiento a largo plazo, pueden modificarse genéticamente líneas celulares y sistemas huésped que expresen de forma estable el producto génico.

Las células huésped y en particular las líneas celulares producidas utilizando los métodos descritos en la presente memoria pueden ser particularmente útiles en la identificación y evaluación de los compuestos que modulan la actividad de una proteína de KLK15.

Pueden expresarse también las proteínas de la invención en animales transgénicos no humanos incluyendo pero sin limitarse a ratones, ratas, conejos, cobayas, lechones, cabras, ovejas, cerdos, primates no humanos (p. ej. mandriles, monos y chimpancés) [véase Hammer *et al.* (*Nature* 315:680-683, 1985), Palmiter *et al.* (*Science* 222:809-814, 1983), Brinster *et al.* (*Proc. Natl. Acad. Sci. USA* 82:4438-4442, 1985), Palmiter y Brinster (*Cell* 41:343-345, 1985) y patente US nº 4.736.866]. Los procedimientos conocidos en la materia pueden utilizarse para introducir una molécula de ácido nucleico de la invención que codifica una proteína de KLK15 en animales para producir las líneas de base de animales transgénicos. Dichos procedimientos incluyen la microinyección pronuclear, la transferencia génica mediada por retrovirus en líneas germinales, el direccionamiento génico en células madre embrionarias, la electroporación del embrión y la transferencia génica mediada por espermatozoides.

La presente invención contempla un animal transgénico que lleva el gen KLK15 en todas sus células y los animales que llevan el transgén en algunas pero no en todas sus células. El transgén puede estar integrado como un único transgén o en concatámeros. El transgén puede introducirse de manera selectiva y activarse en tipos de células específicas (véase por ejemplo, Lasko *et al.*, 1992 *Proc. Natl. Acad. Sci. USA* 89: 6236). El transgén puede integrarse en la secuencia cromosómica del gen endógeno por direccionamiento génico. El transgén puede introducirse selectivamente en un tipo de célula determinado inactivando el gen endógeno en este tipo celular (véase Gu *et al. Science* 265: 103-106).

La expresión de una proteína de KLK15 recombinante en un animal transgénico puede analizarse utilizando técnicas normalizadas. La identificación inicial puede realizarse mediante análisis por transferencia Southern o por métodos PCR para analizar si se ha integrado el transgén. Puede evaluarse también el nivel de expresión de ARNm en los tejidos de animales transgénicos utilizando las técnicas que incluyen el análisis por transferencia Northern de muestras de tejido, la hibridación *in situ* y RT-PCR. El tejido puede evaluarse también por inmunocitoquímica utilizando anticuerpos contra la proteína de KLK15.

Pueden prepararse también proteínas de la invención mediante síntesis química utilizando técnicas bien conocidas en la química de las proteínas tales como la síntesis en fase sólida (Merrifield, 1964, *J. Am. Chem. Assoc.* 85:2149-2154) o la síntesis en solución homogénea (Houbenweyl, 1987, *Methods of Organic Chemistry*, ed. E. Wansch, vol. 15 I y II, Thieme, Stuttgart).

Las proteínas de fusión con terminal N o con terminal C que comprenden la proteína de KLK15 de la invención conjugada con otras moléculas, tales como proteínas, pueden prepararse fusionando, por técnicas recombinantes, el terminal N o el terminal C de una proteína de KLK15 y la secuencia de una proteína o proteína marcadora seleccionadas con una función biológica deseada. Las proteínas de fusión resultantes contienen una proteína de KLK15 fusionada a la proteína o proteína marcadora seleccionada como se describe en la presente memoria. Ejemplos de proteínas que pueden utilizarse para preparar las proteínas de fusión incluyen las inmunoglobulinas, glutatión-S-transferasa (GST), hemoglobulina (HA) y myc truncada.

3. Anticuerpos

Las proteínas de KLK15 de la invención pueden utilizarse para preparar anticuerpos específicos para las proteínas. Pueden prepararse anticuerpos que se unen a un epítipo distinto en una zona no conservada de la proteína. Una zona no conservada de la proteína es la que no presenta homología de secuencia sustancial con otras proteínas. Puede también utilizarse una zona de una zona conservada tal como un dominio bien caracterizado para preparar un anticuerpo contra una zona conservada de una proteína de KLK15. Pueden también producirse anticuerpos con especificidad para una proteína de KLK15 a partir de las proteínas de fusión creadas expresando proteínas de fusión en bacterias como se describe en la presente memoria.

La invención puede emplear anticuerpos íntegros monoclonales o policlonales y fragmentos inmunológicamente activos (p. ej. un fragmento Fab, (Fab)₂ o fragmentos del banco de expresión de Fab y fragmentos que se unen al epítipo de los mismos), una cadena pesada de anticuerpo y una cadena ligera de anticuerpo, una molécula Fv de una sola cadena modificada genéticamente (Ladner *et al.*, patente US nº 4.946.778), anticuerpo humanizado o un anticuerpo híbrido, por ejemplo, un anticuerpo que contiene la especificidad de un unión de un anticuerpo murino, pero en el que las partes restantes son de origen humano. Pueden prepararse anticuerpos que incluyen anticuerpos monoclonales y policlonales, fragmentos e híbridos, utilizando los métodos conocidos por los expertos en la materia.

4. Aplicaciones de las moléculas de ácido nucleico, proteínas de KLK15 y anticuerpos de la invención

Las moléculas de ácido nucleico, las proteínas de KLK15 y los anticuerpos de la invención pueden utilizarse en la evaluación del pronóstico y del diagnóstico de trastornos que involucran una proteína de KLK15 (p. ej. cáncer o trastornos del tiroides) y en la identificación de pacientes con una predisposición a dichos trastornos (apartado 4.1.1 y 4.1.2).

La presente invención proporciona un kit de ensayo para diagnosticar una enfermedad asociada a una proteína de la invención determinando la presencia de una molécula de ácido nucleico o de una proteína de la invención.

El kit de ensayo puede ser utilizado por un método para detectar la expresión del marcador KLK15 de cáncer en un paciente que comprende:

- (a) tomar una muestra procedente de un paciente; y
- (b) detectar en la muestra una secuencia de ácido nucleico que codifica KLK15 o un producto de la proteína codificado por una secuencia de ácido nucleico con KLK15.

Las moléculas de ácido nucleico, las proteínas de KLK15 y los anticuerpos de la invención pueden utilizarse en el diagnóstico y estadificación del cáncer, en particular del cáncer de próstata. El aumento de las concentraciones de las proteínas de KLK15 se asocia con formas no agresivas del cáncer de próstata y puede ser un indicador de pronóstico escaso.

Los métodos para detectar las moléculas de ácido nucleico y las proteínas de KLK15 de la invención, pueden utilizarse para controlar los trastornos que involucran una proteína de KLK15 detectando las proteínas de KLK15 y las moléculas de ácido nucleico que codifican las proteínas de KLK15. Las aplicaciones de la presente invención incluyen también métodos para la identificación de compuestos que modulan la actividad biológica de las proteínas de KLK15 (apartado 4.2). Los compuestos, anticuerpos, etc. pueden utilizarse para el tratamiento de trastornos que involucran una proteína de KLK15 (sección 4.3). Sería asimismo evidente para un experto en la materia que los métodos descritos en la presente memoria puedan utilizarse para estudiar la expresión del desarrollo de las proteínas de KLK15 y, por consiguiente, proporcionará más comprensión en la función de las proteínas de KLK15.

4.1 Métodos de diagnóstico

Pueden emplearse varios métodos para la evaluación del diagnóstico y del pronóstico de trastornos que involucran una proteína de KLK15 y para la identificación de pacientes con una predisposición a dichos trastornos. Dichos métodos, por ejemplo, pueden utilizar moléculas de ácido nucleico de la invención y fragmentos de las mismas y anticuerpos dirigidos contra las proteínas de KLK15, incluyendo fragmentos peptídicos. En particular, pueden utilizarse ácidos nucleicos y anticuerpos, por ejemplo, para: (1) la detección de la presencia de mutaciones de KLK15 o para la detección de la sobreexpresión o infraexpresión del ARNm de KLK15 en comparación con un estado sin trastorno o para la detección cualitativa o cuantitativa de formas alternativamente cortadas y empalmadas de los transcritos de KLK15 que pueden correlacionarse con determinadas enfermedades o la susceptibilidad hacia dichas dolencias; y (2) la detección de una sobreabundancia o una infraabundancia de proteínas de KLK15 en comparación con un estado sin trastorno o la presencia de una proteína de KLK15 modificada (p. ej., menos de la longitud total) que se correlaciona con un estado del trastorno o una evolución hacia un estado del trastorno.

Los métodos descritos en la presente memoria pueden utilizarse para evaluar la probabilidad de la presencia de células malignas o premalignas, por ejemplo, en un grupo de células recién eliminadas de un huésped. Dichos métodos pueden utilizarse para detectar tumores, cuantificar su crecimiento y ayudar al diagnóstico y pronóstico de la enfermedad. Los métodos pueden utilizarse para detectar la presencia de metástasis de cáncer, así como para confirmar la ausencia o eliminación de todo el tejido tumoral tras la cirugía, quimioterapia del cáncer y/o la terapia de radiación. Pueden además utilizarse para controlar la quimioterapia del cáncer y la reaparición del tumor.

Pueden realizarse los métodos descritos en la presente memoria utilizando kits de diagnóstico preenvasados que comprenden por lo menos un ácido nucleico o un anticuerpo específicos para KLK15 descritos en la presente memoria que pueden utilizarse convenientemente, p. ej., en instalaciones clínicas, para observar y diagnosticar pacientes y para observar e identificar aquellos individuos que presentan una predisposición a desarrollar un trastorno.

Las técnicas de detección basadas en ácidos nucleicos se describen, a continuación, en el apartado 4.1.1. Las técnicas de detección de péptidos se describen, a continuación, en el apartado 4.1.2. Las muestras que pueden analizarse utilizando los métodos de la invención incluyen las que son conocidas o se sospecha que expresan a KLK15 o contienen proteínas de KLK15. Las muestras pueden proceder de un paciente o de un cultivo celular e incluyen pero no se limitan a fluidos biológicos, extractos tisulares, células recién recogidas y lisados de células que han sido incubados en cultivos celulares.

Los oligonucleótidos o fragmentos mayores procedentes de algunos de los ácidos nucleicos de la invención pueden utilizarse como dianas en una micromatriz. La micromatriz pueden utilizarse para controlar simultáneamente los niveles de expresión de grandes cantidades de genes y para identificar variantes, mutaciones y polimorfismos genéticos. La información de la micromatriz puede utilizarse para determinar la función génica, para comprender las bases genéticas de una dolencia, para diagnosticar un trastorno y para desarrollar y controlar las actividades de agentes terapéuticos.

La preparación, utilización y análisis de micromatrices son bien conocidas por un experto en la materia. (Véase, por ejemplo, Brennan, T. M. *et al.* (1995) patente US n° 5.474.796; Schena, *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler *et al.* (1995), solicitud PCT WO 95/251116; Shalon, D. *et al.* (1995) solicitud PCT WO 95/35505; Heller, R. A. *et al.* (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-2155; y Heller M. J. *et al.* (1997) patente US n° 5.605.662).

4.1.1 Métodos para detectar moléculas de ácido nucleico de la invención

Las moléculas de ácido nucleico de la invención permiten a los expertos en la materia construir sondas de nucleótido para su utilización en la detección de secuencias de ácido nucleico de la invención en muestras. Las sondas adecuadas incluyen moléculas de ácido nucleico basadas en secuencias de ácido nucleico que codifican por lo menos 5 aminoácidos sucesivos de las zonas de la proteína de KLK15, preferentemente comprenden 15 a 30 nucleótidos (véanse las SEC. ID n°: 47 a 50). Una sonda de nucleótido puede estar marcada con una sustancia detectable tal como un marcador radioactivo que proporciona una señal adecuada y presenta suficiente vida media tal como ³²P, ³H, ¹⁴C o similares. Otras sustancias detectables que pueden utilizarse incluyen los antígenos que son reconocidos por un anticuerpo específico marcado, compuestos fluorescentes, enzimas, anticuerpos específicos para un antígeno marcado y compuestos luminiscentes. Un marcador apropiado puede seleccionarse con respecto a la velocidad de hibridación y a la unión de la sonda al nucleótido que debe detectarse y a la cantidad de nucleótido disponible para hibridación. Las sondas marcadas pueden hibridarse a ácidos nucleicos en soportes sólidos tales como fibras de nitrocelulosa o

membranas de nilón como se describe generalmente en Sambrook *et al.*, 1989, *Molecular Cloning, A Laboratory Manual* (2ª ed.). Pueden utilizarse sondas de ácido nucleico para detectar genes, preferentemente en células humanas, que codifican proteínas de KLK15. Pueden utilizarse sondas de nucleótido también en el diagnóstico de trastornos que involucran una proteína de KLK15; en el control de la evolución de dichos trastornos; o en el control de un tratamiento terapéutico.

La sonda puede utilizarse en técnicas de hibridación para detectar los genes que codifican proteínas de KLK15. La técnica conlleva generalmente poner en contacto e incubar ácidos nucleicos (por ejemplo molécula de ADN recombinante, genes clonados) obtenidos de una muestra procedente de un paciente o de otra fuente celular con una sonda de la presente invención en condiciones favorables para la hibridación específica de las sondas con secuencias complementarias en los ácidos nucleicos. Después de la incubación, se eliminan los ácidos nucleicos no hibridados y se detecta la presencia de los ácidos nucleicos que se han hibridado con la sonda si existen.

La detección de las moléculas de ácido nucleico de la invención puede implicar la ampliación de secuencias génicas específicas que utilizan un método de ampliación tal como PCR, seguido del análisis de las moléculas ampliadas utilizando técnicas conocidas por los expertos en la materia. Los expertos en la materia pueden diseñar de forma rutinaria cebadores adecuados.

Puede utilizarse ADN genómico en ensayos de hibridación o ampliación de muestras biológicas para detectar anomalías que involucran a la estructura de KLK15, incluyendo mutaciones, inserciones, deleciones y transposiciones cromosómicas puntuales. Por ejemplo, puede utilizarse el secuenciado directo, análisis de polimorfismo conformacional de una sola cadena, análisis heterodoble, electroforesis en gel con gradiente de desnaturalización, escisión química incompatible e hibridación de oligonucleótidos.

Pueden utilizarse técnicas de genotipia conocidas por un experto en la materia para tipificar polimorfismos que están en proximidad íntima con las mutaciones en un gen *klk15*. Pueden utilizarse polimorfismos para identificar individuos en familias que es probable que lleven mutaciones. Si un polimorfismo presenta desequilibrio de enlace con las mutaciones en un gen con KLK15, puede utilizarse también la identificación para los individuos en una población general que es probable que lleve mutaciones. Los polimorfismos que pueden utilizarse incluyen polimorfismos con longitud del fragmento de restricción (RFLP), polimorfismos de una sola base y polimorfismos sencillos con repetición de la secuencia (SSLP).

Puede utilizarse una sonda de la invención para identificar directamente los RFPL. Puede utilizarse además una sonda o cebador de la invención para aislar clones genómicos tales como YAC, BAC, PAC, cósmidos, fago o plásmidos. El ADN en los clones puede identificarse para los SSLP utilizando procedimientos de hibridación o de secuenciado.

Pueden utilizarse las técnicas de hibridación y ampliación descritas en la presente memoria en aspectos de análisis cualitativo y cuantitativo de la expresión de *klk15*. Por ejemplo, puede aislarse ARN de un tipo de célula o tejido conocido para expresar *klk15* y analizarse utilizando la hibridación (p. ej. análisis de Northern normalizado) o técnicas de PCR mencionadas en la presente memoria. Pueden utilizarse técnicas para detectar diferencias en el tamaño del transcrito que pueden ser debido a corte y empalme alternativo normal o anormal. Pueden utilizarse las técnicas para detectar diferencias cuantitativas entre niveles de transcritos completos y/o alternativamente de corte y empalme detectados en individuos normales en comparación con los individuos que presentan síntomas de un trastorno que involucra una proteína de KLK15.

Pueden utilizarse cebadores y sondas en los métodos *in situ* descritos anteriormente, es decir directamente en las secciones de tejido (fijadas y/o congeladas) del tejido del paciente extraído por biopsias y resecciones.

4.1.2 Métodos para detectar proteínas de KLK15

Pueden utilizarse anticuerpos específicamente reactivos con una proteína de KLK15, o derivados, tales como conjugados enzimáticos o derivados marcados, para detectar proteínas de KLK15 en varias muestras (por ejemplo, materiales biológicos). Pueden utilizarse como reactivos para diagnóstico o pronóstico y pueden utilizarse para detectar anomalías en el nivel de expresión de la proteína de KLK15, o anomalías en la estructura y/o posición temporal, del tejido, celular o subcelular de una proteína de KLK15. Pueden utilizarse anticuerpos para identificar *in vitro* compuestos potencialmente terapéuticos para determinar sus efectos sobre los trastornos que involucran una proteína de KLK15, y otras enfermedades. Puede utilizarse también inmunoanálisis *in vitro* para evaluar o controlar la eficacia de determinadas terapias. Pueden utilizarse también los anticuerpos de la invención *in vitro* para determinar el nivel de expresión de KLK15 en células modificadas genéticamente para producir la proteína de KLK15.

Pueden utilizarse anticuerpos en cualquier inmunoanálisis conocido que se base en la interacción del enlace entre un determinante antigénico de una proteína de KLK15 y los anticuerpos. Ejemplos de dichos análisis son el radioinmunoanálisis, el inmunoanálisis enzimático (p. ej. ELISA), la inmunofluorescencia, la inmunoprecipitación, la aglutinación en látex, la hemoaglutinación y las pruebas histoquímicas. Pueden utilizarse anticuerpos para detectar y cuantificar proteínas de KLK15 en una muestra a fin de determinar su función en determinados episodios celulares o estados patológicos y para diagnosticar y tratar dichos estados patológicos.

En particular, pueden utilizarse los anticuerpos de la invención en análisis inmunohistoquímicos, por ejemplo, en el nivel celular y subsubcelular, para detectar una proteína de KLK15, para localizarla en determinadas células y tejidos y para especificar posiciones subcelulares y para cuantificar el nivel de expresión.

5 Pueden utilizarse técnicas histoquímicas conocidas en la materia para localizar antígenos que utilizan microscopía óptica y electrónica para detectar una proteína de KLK15. Generalmente, puede marcarse un anticuerpo de la invención con una sustancia detectable y una proteína relacionada con KLK15 puede localizarse en los tejidos y células basándose en la presencia de la sustancia detectable. Ejemplos de sustancias detectables incluyen, pero no se limitan a, los siguientes radioisótopos (p. ej., ^3H , ^{14}C , ^{35}S , ^{125}I , ^{131}I), marcadores fluorescentes (p. ej., FITC, rodamina, lantánidos fosforescentes), marcadores luminiscentes tal como luminol; marcadores enzimáticos (p. ej. peroxidasa de rábano picante, β -galactosidasa, luciferasa, fosfatasa alcalina, acetilcolinesterasa), grupos biotínico (que pueden detectarse mediante avidina marcada p. ej., estreptavidina que contiene un marcador fluorescente o actividad enzimática que puede detectarse por métodos ópticos o calorimétricos), epítomos de polipéptido predeterminados reconocidos por un indicador secundario (p. ej., secuencias par cremallera de leucina, secuencias de enlace para anticuerpos secundarios, 10 dominios de enlace metálico, etiquetas de epítomo). En algunas formas de realización, los marcadores están unidos mediante brazos espaciadores de varias longitudes para reducir el impedimento estérico potencial. Los anticuerpos pueden también acoplarse a sustancias de densidad electrónica, tal como ferritina u oro coloidal, que se observan fácilmente al microscopio electrónico.

20 El anticuerpo o la muestra pueden inmovilizarse sobre un portador o soporte sólido que sea capaz de inmovilizar las células, anticuerpos, etc. Por ejemplo, el portador o soporte puede ser nitrocelulosa, o vidrio, poliacrilamidas, gabbros y magnetita. El material del soporte puede presentar cualquier configuración posible incluyendo la esférica (p. ej. perlas), cilíndrica (p. ej. superficie interna de un tubo de ensayo o pocillos, o la superficie externa de una varilla) o plana (p. ej. lámina, tira de prueba). Pueden emplearse también métodos indirectos en los que la reacción antígeno-anticuerpo primaria se amplía mediante la introducción de un segundo anticuerpo, que presenta especificidad para el anticuerpo reactivo frente a la proteína de KLK15. A título de ejemplo, si el anticuerpo que presenta especificidad a la proteína de KLK15 es un anticuerpo de IgG de conejo, el segundo anticuerpo puede ser gamma-globulina anticonejo de cabra marcada con una sustancia detectable como se describe en la presente memoria.

30 Cuando se utiliza un marcador radioactivo como sustancia detectable, puede localizarse una proteína de KLK15 por radioautografía. Los resultados de la radioautografía pueden cuantificarse determinando la densidad de partículas en los radioautógrafos por varios métodos ópticos, o mediante recuento de los granos.

35 En una forma de realización, la invención contempla un kit de ensayo para controlar la evolución del cáncer (p. ej. cáncer de próstata) en un individuo, que comprende:

- (a) poner en contacto una cantidad de un anticuerpo que se une a la proteína de KLK15, con una muestra del individuo de modo que formen un complejo binario que comprende el anticuerpo y la proteína de KLK15 en la muestra;
- (b) determinar o detectar la presencia o cantidad de formación de complejo en la muestra;
- (c) repetir las etapas (a) y (b) en un punto posterior en el tiempo; y
- (d) comparar el resultado de la etapa (b) con el resultado de la etapa (c), en el que una diferencia en la cantidad de formación de complejo es indicativa de la evolución del cáncer en dicho individuo.

La cantidad de complejos puede compararse también con un valor representativo de la cantidad de complejos de un individuo no en riesgo de cáncer ni afectado por éste (p. ej. cáncer de próstata).

4.2 Métodos para identificar o evaluar sustancias/compuestos

Los métodos descritos en la presente memoria se diseñan para identificar sustancias que modulen la actividad biológica de una proteína de KLK15 incluyendo las sustancias que se unen a proteínas de KLK15, o unirse a otras proteínas que interactúan con una proteína de KLK15, a compuestos que interfieren con la interacción de una proteína de KLK15 o aumentan ésta y sustancias que se unen a la proteína de KLK15 u otras proteínas que interactúan con una proteína de KLK15. Los métodos se utilizan también para identificar compuestos que se unen a las secuencias reguladoras de KLK15.

60 Las sustancias y compuestos identificados utilizando los métodos de la invención incluyen pero no se limitan a los péptidos tales como los péptidos solubles que incluyen péptidos de fusión con cola de Ig, elementos de bancos de péptidos aleatorios y bancos moleculares derivados de la química combinatoria de aminoácidos en configuración D y/o L, fosfopéptidos (incluyendo los elementos de azar o parcialmente degenerados, bancos de fosfopéptidos dirigidos), anticuerpos [p. ej., policlonales, monoclonales, humanizados, antiidiotípico, híbrido, anticuerpos de una sola cadena, fragmentos (p. ej., Fab, F(ab)₂ y fragmento del banco de expresión de Fab y fragmentos de estos de unión al epítomo)], y pequeñas moléculas orgánicas o inorgánicas. La sustancia o compuesto puede ser un compuesto fisiológico endógeno o puede ser un compuesto natural o sintético.

Las sustancias que modulan una proteína de KLK15 pueden identificarse basándose en su capacidad para unirse a una proteína de KLK15. Por consiguiente la invención también proporciona métodos para identificar sustancias que se unen a la proteína de KLK15. Las sustancias identificadas que utilizan los métodos de la invención pueden aislarse, clonarse y secuenciarse utilizando técnicas convencionales. Una sustancia que se asocia a un polipéptido de la invención puede ser un agonista o antagonista de la actividad biológica o inmunológica de un polipéptido de la invención.

El término “agonista”, se refiere a una molécula que aumenta la cantidad de actividad de la proteína o prolonga su duración. El término “antagonista” se refiere a una molécula que disminuye la actividad biológica o inmunológica de la proteína. Pueden incluirse agonistas y antagonistas, proteínas, ácidos nucleicos, carbohidratos o pueden asociarse otras moléculas con una proteína de la invención.

Sustancias que se unen con una proteína de KLK15 pueden identificarse haciendo reaccionar una proteína de KLK15 con una sustancia de la prueba que se une potencialmente a una proteína de KLK15, en condiciones que permitan la formación de los complejos proteicos sustancia-KLK15 y eliminar y/o detectar los complejos. Los complejos pueden determinarse analizando la sustancia libre o la proteína de KLK15 no acomplejada en los complejos sustancia-proteína de KLK15. Las condiciones que permiten la formación de complejos sustancia-proteína de KLK15 pueden seleccionarse con respecto a factores tales como la naturaleza y cantidades de la sustancia y de la proteína.

El complejo sustancia-proteína, la sustancia libre o las proteínas no acomplejadas pueden aislarse mediante técnicas de aislamiento convencional, por ejemplo, precipitación por salado, cromatografía, electroforesis, filtración en gel, fraccionamiento, absorción, electroforesis en gel de poliácridamida, aglutinación o combinaciones de los mismos. Para facilitar el análisis de los componentes, puede utilizarse el anticuerpo contra la proteína de KLK15 o la sustancia, o la proteína de KLK15 marcada, o una sustancia marcada. Los anticuerpos, proteínas o sustancias pueden marcarse con una sustancia detectable como se describió anteriormente.

Una proteína de KLK15 o la sustancia utilizada en el método de la invención pueden insolubilizarse. Por ejemplo, una proteína de KLK15 o la sustancia pueden estar unidas a un portador adecuado tal como agarosa, celulosa, dextrano, Sephadex, sepharosa, carboximetilcelulosa poliestireno, papel de filtro, resina de intercambio iónico, película de plástico, tubo de plástico, copolímero poliamina-éter metilvinílico-ácido maleico, copolímero de aminoácidos, copolímero de etileno-ácido maleico, nilón, seda, etc. El portador puede estar en forma, por ejemplo, de tubo, lámina de prueba, perlas, disco, esfera, etc. La proteína o sustancia insolubilizada puede prepararse haciendo reaccionar el material con un portador insoluble adecuado utilizando métodos químicos o físicos conocidos, por ejemplo, acoplamiento con bromuro de cianógeno.

La invención contempla también un método para evaluar un compuesto por su capacidad para modular la actividad biológica de una proteína de KLK15 de la invención, analizando un agonista o antagonista (es decir potenciador o inhibidor) del enlace de una proteína de KLK15 con una sustancia que se une con una proteína de KLK15. El método básico para evaluar si un compuesto es un agonista o antagonista del enlace de una proteína de KLK15 y una sustancia que se une a la proteína, consiste en preparar una mezcla de reacción que contiene la proteína de KLK15 y la sustancia en condiciones que permitan la formación de complejos sustancia-proteína de KLK15, en presencia de un compuesto de ensayo. El compuesto de ensayo puede añadirse inicialmente a la mezcla, o puede añadirse después de la adición de la proteína de KLK15 y la sustancia. También se preparan mezclas de reacción de referencia sin el compuesto de ensayo o con un placebo. Se detecta la formación de complejos y la formación de complejos en la reacción de control pero no en la mezcla de reacción indica que el compuesto de la prueba interfiere con la interacción de la proteína de KLK15 y la sustancia. Las reacciones pueden realizarse en fase líquida o la proteína de KLK15, la sustancia o el compuesto de la prueba pueden estar inmovilizados como se describe en la presente memoria. La capacidad de un compuesto para modular la actividad biológica de una proteína de KLK15 de la invención puede analizarse determinando los efectos biológicos sobre las células.

Debe entenderse que los agonistas y antagonistas, es decir los inhibidores y los potenciadores que pueden analizarse utilizando los métodos de la invención pueden actuar en una o más de las secuencias de enlace sobre la proteína o la sustancia incluyendo las secuencias de enlace del agonista, las secuencias de enlace del antagonista competitivo, las secuencias de enlace del antagonista no competitivo o las secuencias alostéricas.

La invención también permite identificar antagonistas que inhiben los efectos de un agonista de la interacción de la proteína de KLK15 con una sustancia que es capaz de unirse a la proteína de KLK15. Por esta razón, la invención puede utilizarse para analizar un compuesto que compite por la misma secuencia de enlace de una proteína de KLK15.

La invención contempla también métodos para identificar compuestos que se unen a proteínas que interactúan con una proteína de KLK15. Las interacciones proteína-proteína pueden ser identificadas utilizando métodos convencionales tal como la coimmunoprecipitación, reticulación y purificación conjunta mediante gradientes o columnas cromatográficas. Pueden emplearse también métodos que produzcan la identificación simultánea de genes que codifican proteínas que interactúan con una proteína de KLK15. Estos métodos incluyen el sondado de bancos de expresión con la proteína de KLK15 marcada.

Pueden utilizarse también sistemas de dos híbridos para detectar interacciones de la proteína *in vivo*. Generalmente, se construyen plásmidos que codifican dos proteínas híbridas. Una primera proteína híbrida consta del dominio

de enlace al ADN de una proteína activadora de transcripción fusionada a una proteína de KLK15 y la segunda proteína híbrida consta del dominio del activador de la proteína del activador de transcripción fusionado a una proteína desconocida codificada por un ADNc que se ha recombinado en el plásmido con la parte de un banco de ADNc. Los plásmidos se transforman en una cepa de levadura (p. ej. *S. cerevisiae*) que contiene un gen indicador (p. ej. lacZ, luciferasa, fosfatasa alcalina, peroxidasa de rábano picante) cuya zona reguladora contiene la secuencia de enlace del activador de transcripción. Las proteínas híbridas solas no pueden activar la transcripción del gen indicador. Sin embargo, la interacción de las dos proteínas híbridas reconstituye la proteína activadora operativa y produce la expresión del gen indicador, que se detecta mediante un análisis por el producto del gen indicador.

Debe apreciarse que las proteínas de fusión pueden utilizarse en los métodos descritos anteriormente. En particular, las proteínas de KLK15 fusionadas a una glutatión-S-transferasa pueden utilizarse en los métodos.

Un modulador de una proteína de KLK15 de la invención puede también identificarse basándose en su capacidad para inhibir o potenciar la actividad catalítica de la proteína.

Los reactivos adecuados para aplicar a los métodos de la invención destinados a evaluar los compuestos que modulan una proteína de KLK15 pueden estar empaquetados en kits convenientes proporcionando los materiales necesarios empaquetados en recipientes adecuados. Los kits pueden incluir también soportes adecuados útiles para realizar los métodos de la invención.

4.3. Composiciones y tratamientos

Las proteínas de la invención, las sustancias o los compuestos identificados mediante los métodos descritos en la presente memoria, los antibióticos y las moléculas de ácido nucleico de la invención pueden utilizarse para modular la actividad biológica de una proteína de KLK15 y pueden utilizarse en el tratamiento de enfermedades tal como el cáncer (particularmente el cáncer de tiroides, de próstata, de colon, de riñón y testicular) y trastornos del tiroides en un paciente.

Por consiguiente, pueden formularse sustancias, anticuerpos, péptidos y compuestos en composiciones farmacéuticas para la administración a pacientes de forma biológicamente compatible adecuada para la administración *in vivo*. “Forma biológicamente compatible adecuada para su administración *in vivo*” significa una forma de la sustancia activa que debe administrarse en la que algunos efectos tóxicos están sobrevalorados por los efectos terapéuticos. Las sustancias activas pueden administrarse a organismos vivos incluyendo los seres humanos y los animales. La administración de una cantidad terapéuticamente activa de una composición farmacéutica de la presente invención está definida como una cantidad eficaz, a las dosis y durante periodos de tiempo necesarios para alcanzar todo resultado deseado. Por ejemplo, una cantidad terapéuticamente activa de una sustancia puede variar según factores tales como la enfermedad, edad, sexo y peso del individuo, y la capacidad del anticuerpo para producir una respuesta deseada en el individuo. Los regímenes de dosificación puede ajustarse para proporcionar la respuesta terapéutica óptima. Por ejemplo, pueden administrarse a diario varias dosis inhibidas o la dosis puede reducirse proporcionalmente como se indica por exigencias de la solución terapéutica.

La sustancia activa puede administrarse de manera conveniente tal como mediante inyección (subcutánea, intravenosa, etc.), administración oral, inhalación, aplicación transdérmica o administración rectal. Dependiendo de la vía de administración, la sustancia activa puede estar recubierta de un material que proteja la sustancia de la acción de las enzimas, ácidos y otras condiciones naturales que puedan inactivar la sustancia.

Las composiciones descritas en la presente memoria pueden prepararse por métodos conocidos por sí mismos para la preparación de composiciones farmacéuticamente aceptables que pueden administrarse a pacientes, de modo que una cantidad de la sustancia activa se combina en una mezcla con un vehículo farmacéuticamente aceptable. Los vehículos adecuados se describen, por ejemplo, en Remington's Pharmaceutical Sciences (Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pa., USA 1985). Basándose en esto, las composiciones incluyen, aunque no exclusivamente, soluciones de las sustancias activas junto con uno o más vehículos o diluyentes farmacéuticamente aceptables y contenidos en soluciones tamponadas con un pH adecuado e isoosmóticas con los fluidos fisiológicos.

Las composiciones están indicadas como agentes terapéuticos bien solos o junto con otros agentes terapéuticos u otras formas de tratamiento (p. ej. quimioterapia o radioterapia). Por ejemplo, pueden utilizarse composiciones en combinación con agentes antiproliferantes, agentes antimicrobianos, agentes inmunoestimulantes o antiinflamatorios. En particular, los compuestos pueden utilizarse en combinación con agentes antivíricos y/o antiproliferantes. Las composiciones de la invención pueden administrarse simultáneamente, por separado o sucesivamente con otros agentes terapéuticos o terapias.

Los vectores procedentes de retrovirus, adenovirus, herpes virus o vacunas o procedentes de varios plásmidos bacterianos, pueden utilizarse para administrar moléculas de ácido nucleico a un órgano, tejido o población celular objetivos. Pueden utilizarse métodos bien conocidos por los expertos en la materia para construir vectores recombinantes que expresen moléculas de ácido nucleico con cadena complementaria de la invención. (Véanse, por ejemplo, las técnicas descritas en Sambrook *et al.* (*supra*) y Ausubel *et al.* (*supra*)).

Las moléculas de ácido nucleico que comprenden las secuencias completas de ADNc y/o sus elementos reguladores permiten utilizar a un experto en la materia secuencias que codifican una proteína de la invención como herramienta de investigación en la regulación de la cadena transcrita (Yousoufian H. y H. F. Lodish 1993 *Mol. Cell Biol.* 13:98-104) o de la cadena complementaria (Eguchi *et al.* (1991) *Annu. Rev. Biochem.* 60:631-652) de la función génica.

5 Dicha tecnología es bien conocida en la materia y los oligómeros de cadena transcrita o complementaria o fragmentos mayores, pueden diseñarse en varias posiciones a lo largo de las zonas de codificación o de referencia.

Los genes que codifican una proteína de la invención pueden desviarse transfectando una célula o tejido con vectores que expresan grandes niveles de un fragmento deseado de codificación de KLK15. Dichos montajes pueden inundar las células con secuencias de cadena transcrita o complementaria no traducibles. Aún en ausencia de integración en el ADN, dichos vectores pueden continuar para transcribir moléculas de ARN hasta que todas las copias estén deshabilitadas por nucleasas endógenas.

10

Pueden obtenerse modificaciones de la expresión génica diseñando moléculas de cadena complementaria, ADN, ARN o APN, en las zonas reguladoras de un gen que codifica una proteína de la invención, es decir, los activadores, potenciadores e intrones. Preferentemente, los oligonucleótidos proceden de la secuencia de iniciación de la transcripción, p. ej. Entre las zonas -10 y +10 de la secuencia principal. Pueden también diseñarse moléculas de cadena complementaria de modo que bloqueen la traducción del ARNm impidiendo la unión del transcrito a los ribosomas. La inhibición también puede conseguirse utilizando la metodología de emparejamiento de bases de "triple hélice". El emparejamiento de triple hélice compromete la capacidad de la doble hélice de abrir suficientemente la unión de las polimerasas, los factores de transcripción o las moléculas reguladoras. Los avances terapéuticos que utilizan ADN triple fueron estudiados por Gee J. E. *et al.* (en: Huber B. E. y B. I. Carr (1994) *Molecular and Immunologic Approaches*, Futura Publishing Co, Mt Kisco N.Y.).

15 20

Los ribozimas son moléculas de ARN enzimático que catalizan la escisión específica del ARN. Los ribozimas actúan por hibridación específica de la secuencia de la molécula de ribozima al ARN diana complementario, seguido de escisión endonucleolítica. La invención contempla por lo tanto moléculas de ribozima con motivo de cabeza de martillo modificadas genéticamente que pueden catalizar de manera específica y eficaz la escisión endonucleolítica de las secuencias que codifican una proteína de la invención.

25 30

Las secuencias específicas para la escisión del ribozima en cualquier ARN potencial diana pueden identificarse inicialmente explorando las secuencias de adición del ribozima en la molécula diana que incluyen las secuencias siguientes, GUA, GUU y GUC. Una vez identificadas las secuencias, las secuencias del ARN corto de entre 15 y 20 ribonucleótidos correspondientes a la zona del gen diana que contiene la secuencia de escisión pueden evaluarse las características estructurales secundarias que pueden hacer inoperable el oligonucleótido. La idoneidad de las dianas experimentales puede determinarse también probando la accesibilidad a la hibridación con oligonucleótidos complementarios que utilizan ensayos de protección de la ribonucleasa.

35

Los métodos para introducir vectores en células o tejidos incluyen aquellos métodos expuestos en la presente memoria y que son adecuados para la terapia *in vivo*, *in vitro* y *ex vivo*. En la terapia *ex vivo*, los vectores pueden introducirse en células madre extraídas de un paciente y propagadas clonalmente para el trasplante autólogo en el mismo paciente (véase las patentes US nº 5.399.493 y nº 5.437.994). La administración por trasfección y por liposomas son bien conocidas en la materia.

40

Un anticuerpo contra una proteína de KLK15 puede estar conjugado con fármacos quimioterapéuticos, toxinas, modificadores de la respuesta inmunológica, agentes hematógenos, enzimas y radioisótopos y utilizarse para la prevención y tratamiento del cáncer (p. ej. cáncer de tiroides, próstata, colon, riñón y testicular). Por ejemplo, un anticuerpo contra una proteína de KLK15 puede conjugarse con restos tóxicos que incluyen pero no se limitan a ricina A, toxina de la difteria, abrina, modeccina, o toxinas bacterianas procedentes de *Pseudomonas* o *Shigella*. Se ha descrito que las toxinas y sus derivados forman conjugados con anticuerpos específicos para determinados tejidos diana, tales como células cancerosas o tumorales a fin de obtener toxicidad celular específicamente dirigida (Moolten F. L. *et al.*, *Immun. Rev.* 62:47-72, 1982 y Bernhard, M. I. *Cancer Res.* 43:4420, 1983).

45 50

Pueden prepararse conjugados por los medios habituales conocidos en la materia. Un número de agentes de enlace bifuncionales (p. ej. enlazadores heterobifuncionales tales como N-succinimidil-3-(2-piridilditio)propionato) están disponibles en el mercado en Pierce Chemical Company, Rockford, III.

55

La administración de anticuerpos o inmunotoxinas para uso terapéutico puede ser por vía intravenosa, si bien con formulación apropiada otras vías de administración tales como la administración intraperitoneal, oral o transdérmica también pueden utilizarse.

60

Una proteína de KLK15 puede conjugarse con fármacos quimioterapéuticos, toxinas, modificadores de respuesta inmunológica, enzimas y radioisótopos utilizando métodos conocidos en la materia.

La invención proporciona también métodos inmunoterapéuticos para prevenir o reducir la gravedad de un cáncer. Los signos o síntomas clínicos del cáncer en un paciente son indicativos de un efecto beneficioso en el paciente debido a la estimulación de la respuesta inmunitaria del paciente contra el cáncer. La estimulación de una respuesta inmunitaria se refiere a la producción de una respuesta inmunitaria o la potenciación de la actividad de células inmunoefectoras.

65

5 en respuesta a la administración de una preparación de vacuna de la invención. La prevención de un cáncer puede estar indicada por un aumento del tiempo antes de la aparición del cáncer en un paciente que está predispuesto a desarrollar el cáncer debido por ejemplo a una disposición genética o a la exposición a un agente cancerígeno. La reducción de la gravedad de un cáncer puede estar indicada por una disminución en el tamaño o en la velocidad de crecimiento de un tumor.

10 Las vacunas pueden proceder de una proteína KLK, de los péptidos derivados de la misma o de los péptidos sintéticos producidos químicamente o de cualquier combinación de estas moléculas, o de las proteínas de fusión o de los péptidos de las mismas. Las proteínas, péptidos, etc. pueden sintetizarse o prepararse de manera recombinante o si no biológicamente, para comprender una o más secuencias de aminoácido correspondientes a uno o más epítomos de un tumor asociado a la proteína. Los epítomos de un tumor asociado a la proteína puede entenderse que incluyen la posibilidad de que en algunos casos las variaciones de la secuencia de aminoácidos de una proteína o polipéptido natural pueden ser antigénicas y conferir inmunidad protectora contra el cáncer o los efectos antitumorígenos. Las variaciones de la secuencia pueden incluir sin limitación, sustituciones de aminoácidos, ampliaciones, deleciones, 15 truncamientos, interpolaciones y combinaciones de los mismos. Dichas variaciones están comprendidas dentro del alcance de la invención con la condición de que la proteína que las contiene sea inmunógena y los anticuerpos contra dicho polipéptido interaccionen con la proteína de KLK15 natural en una medida suficiente para proporcionar inmunidad protectora y/o actividad antitumorígena cuando se administran en forma de vacuna.

20 Las proteínas, péptidos, etc. pueden incorporarse a vacunas capaces de producir una respuesta inmunitaria utilizando los métodos conocidos en la materia. Las técnicas para potenciar la antigenicidad de las proteínas, péptidos, etc. son conocidas en la materia e incluyen la incorporación a una estructura multimérica, la unión a un portador de la proteína muy inmunógeno, por ejemplo, la hemocianina de la lapa californiana (KLH) o la vacuna contra la difteria y la administración en combinación con adyuvantes o cualquier otro potenciador de respuesta inmunitaria.

25 Las vacunas pueden combinarse con medios fisiológicamente aceptables, incluyendo diluyentes y portadores inmunológicamente aceptables así como adyuvantes empleados habitualmente tales como el adyuvante completo de Freund, saponina, alúmina y similares.

30 Debe apreciarse además que los anticuerpos antiidiotipo contra los anticuerpos para las proteínas de KLK15 descritos en la presente memoria son también útiles como vacunas y pueden formularse igualmente.

35 La administración de una vacuna según la invención, es aplicable generalmente para la prevención o el tratamiento de cánceres incluyendo el cáncer de tiroides, próstata, colon, riñón y testicular.

40 La administración a un paciente de una vacuna según la invención para la prevención y/o tratamiento del cáncer puede tener lugar antes o después de un procedimiento quirúrgico para eliminar el cáncer, antes o después de un procedimiento terapéutico para el tratamiento del cáncer y antes o después de la terapia de radiación para el tratamiento del cáncer y cualquier combinación de los mismos. La inmunoterapia del cáncer según la invención sería el tratamiento preferido para la prevención y/o tratamiento del cáncer, ya que los efectos secundarios involucrados son sustancialmente mínimos en comparación con los demás tratamientos disponibles, p. ej. cirugía, quimioterapia, terapia de radiación. Las vacunas presentan el potencial o capacidad para prevenir el cáncer en pacientes sin cáncer pero que están en riesgo de desarrollar cáncer.

45 La actividad de las proteínas, sustancias, compuestos, anticuerpos, moléculas de ácido nucleico, agentes y composiciones de la invención pueden estar confirmadas en sistemas de modelo experimental. La eficacia terapéutica y la toxicidad pueden determinarse por procedimientos farmacéuticos normalizados en cultivos celulares o con animales experimentales, tal como calculando la ED₅₀ (dosis terapéuticamente eficaz en el 50% de la población) o la LD₅₀ (dosis letal para el 50% de la población) estadísticas. El índice terapéutico es la relación de dosis de efecto terapéutico atóxico y puede expresarse como la relación ED₅₀/LD₅₀. Se prefieren las composiciones farmacéuticas que presentan efectos terapéuticos grandes.

4.4 Otras aplicaciones

55 Las moléculas de ácido nucleico dadas a conocer en la presente memoria pueden utilizarse también en técnicas de biología molecular que todavía no han sido desarrolladas, con la condición de que las nuevas técnicas basadas en las propiedades de secuencias de nucleótidos que son actualmente conocidas, incluyendo pero sin limitarse a dichas propiedades como el código genético del triplete y las interacciones de pares de bases específicas.

60 A fin de estudiar la función de un polipéptido de la invención, pueden desarrollarse células, tejidos y animales no humanos que carecen de la expresión o que carecen parcialmente de expresión de una molécula de ácido nucleico o gen de la invención utilizando vectores de expresión recombinantes de la invención con mutaciones específicas por delección o infección en el gen. Puede utilizarse un vector de expresión recombinante para inactivar o alterar el gen endógeno por recombinación homóloga y crear de este modo una célula, tejido o animal defectuoso.

65 Ningún alelo puede generarse en células, tales como las células madre embrionarias por mutación con delección. Un gen recombinante puede modificarse genéticamente también para que contenga una mutación por inserción que inactive el gen. Dicho montaje puede introducirse a continuación en una célula, tal como una célula madre embrionaria,

mediante una técnica tal como la transfección, electroporación, inyección, etc. Las células que carecen de un gen íntegro pueden identificarse a continuación, por ejemplo mediante transferencia Southern, transferencia Northern o ensayando la expresión del polipéptido codificado que utiliza los métodos descritos en la presente memoria. Dichas células pueden fusionarse a continuación a células madre embrionarias para generar animales transgénicos no humanos carentes de un polipéptido de la invención. La transmisión por línea germinal de la mutación puede conseguirse, por ejemplo, agregando células madre embrionarias con embriones en etapa inicial, tal como embriones con linfocitos B, *in vitro*; transfiriendo los blastocitos resultantes a hembras receptoras y; generando la transmisión de la línea germinal de los híbridos de agregación resultantes. Dicho animal mutante puede utilizarse para definir poblaciones de células específicas, modelos de desarrollo y procesos *in vivo*, dependientes normalmente de la expresión génica.

Un mamífero transgénico no humano puede desarrollar todas las células germinales y células somáticas que contienen un vector de expresión recombinante que inactiva o altera un gen que codifica una proteína de KLK15. El mamífero transgénico no humano (del que todas las células germinales y las células somáticas contienen un vector de expresión recombinante que inactiva o altera un gen que codifica una proteína de KLK15) pueden producir una patología asociada a la proteína de KLK15. Un mamífero transgénico no humano que no expresa o ha alterado (p. ej. reducido) la expresión de una proteína de KLK15 de la invención puede producir una patología asociada a la proteína de KLK15. Una patología por la proteína de KLK15 se refiere a un fenotipo observado para un mutante homocigótico o heterocigótico de la proteína de KLK15.

Un animal transgénico no humano puede ser, por ejemplo, un ratón, rata, conejo, oveja, hámster, perro, gato, cabra o mono.

Un sistema de ensayo de animal transgénico no humano puede proporcionar un sistema de modelo para determinar un agente que reduce o inhibe una patología asociada a una proteína de KLK15, tal como una patología asociada a la proteína de KLK15, que comprende:

- (a) administrar el agente a un animal transgénico no humano; y
- (b) determinar si dicho agente reduce o inhibe la patología (p. ej. la patología asociada a la proteína de KLK15) en el animal transgénico no humano en comparación con un animal transgénico no humano de la etapa (a) al que no se ha administrado el agente.

El agente puede ser útil en el tratamiento y profilaxis de enfermedades, tales como el cáncer, expuestas en la presente memoria. Los agentes pueden incorporarse también a una composición farmacéutica como se describe en la presente memoria. El ejemplo siguiente no limitativo es ilustrativo de la presente invención:

Ejemplo

Materiales y métodos

Identificación del nuevo gen

Se construyó una cartografía contigua para el locus del gen de calicreína humano que se extiende desde el gen KLK1 (centrómeros) al gen KLK14 (telómero) (7, 8, 11, 12, 27). Se identificaron los clones del cromosoma artificial bacteriano (BAC) solapantes que se extienden por esta área mediante el cribado de un banco BAC humano que utiliza varias sondas específicas para el gen radiomarcadas. Se creó un área de ~300 kb de secuencia genómica utilizando diferentes técnicas, como se describió anteriormente (11, 27). Realizando un análisis de restricción de EcoRI, se orientó el locus de calicreína a lo largo de la cartografía de restricción de EcoRI del cromosoma 19q13 disponible en Lawrence Livermore National Laboratory (LLNL). Se identificó a continuación un clon BAC que se extiende más centroméricamente (BC 781134). Los cóntigos de las secuencias genómicas lineales de este clon están disponibles en el LLNL. Inicialmente, estas secuencias de cóntigo se utilizaban para prever la presencia de nuevos genes, utilizando métodos bioinformáticos, como se describió anteriormente (8, 12) y se identificó una nueva serina proteasa supuesta. La secuencia del supuesto gen se verificó a continuación por diferentes métodos incluyendo el secuenciado, la búsqueda en la base de datos EST, el cribado por PCR de tejidos, como se describe a continuación.

Investigación de la etiqueta de la secuencia expresada (EST)

Los exones previstos del supuesto nuevo gen se sometieron a búsqueda de la homología utilizando el algoritmo BLASTN (28) en el servidor de la red de información del National Center for Biotechnology (<http://www.ncbi.nlm.nih.gov/BLAST/>) frente a la base de datos EST humana (dbEST). Los clones con >95% de homología se obtuvieron del consorcio I.M.A.G.E. (29) a través del Research Genetics Inc, Huntsville, AL. Se propagaron los clones, se purificaron como se describió en otra parte (30) y se secuenciaron en ambas direcciones con un secuenciador automático, utilizando cebadores del vector que flanquea la inserción.

Línea celular de cáncer de próstata y experimentos de estimulación hormonal

La línea celular del cáncer de próstata LNCaP se adquirió en la American Type Culture Collection (ATCC), Rockville, MD. Se cultivaron las células en medio RPMI (Gibco BRL, Gaithersburg, MD) enriquecido con glutamina (200

mmoles/l), insulina bovina (10 mg/l), suero bovino fetal (10%), antibióticos y antimicóticos, en frascos de plástico, hasta casi la confluencia. Se tomaron alícuotas a continuación de las células en placas de cultivo tisular de 24 pocillos y se cultivaron a una confluencia de 50%. 24 horas antes de los experimentos, se cambiaron los medios de cultivo en medio exento de rojo de fenol que contenía suero bovino fetal agotado en carbón activo al 10%. Para los experimentos de estimulación, se añadieron en el medio de cultivo varias hormonas esteroides disueltas en etanol al 100% a una concentración final de 10^{-8} . Las células estimuladas con etanol al 100% se incluyeron como referencias. Se cultivaron las células durante 24 horas, a continuación se recogieron para la extracción del ARNm.

Reacción en cadena de polímera de transcriptasa inversa (RT-PCR) para el gen KLK15

Se extrajo el ARN total de la línea celular LNCaP de tejidos de la próstata utilizando el reactivo Trizol (Gibco BRL) siguiendo las instrucciones del fabricante. Se determinó la concentración de ARN por espectrofotometría. Se transcribieron a la inversa 2 μ g de ARN total en la primera cadena de ADNc utilizando el sistema de preamplificación SuperscriptTM (Gibco BRL). El volumen final fue de 20 μ l. Basándose en la información combinada obtenida de la estructura genómica prevista del nuevo gen y en las secuencias EST (véase a continuación), se diseñaron dos cebadores específicos para el gen (KLK15-F1 - SEC. ID n°: 47 y KLK15-R1 - SEC. ID n°: 48) (Tabla 1) y se realizó la PCR en una mezcla de reacción que contenía 1 μ l de ADNc, Tris-HCl 10 mM (pH 8,3), KCl 50 mM, MgCl₂ 1,5 mM, dNTP 200 μ M (trifosfatos de desoxinucleósido), 150 ng de cebadores y 2,5 unidades de HotStarTM ADN polimerasa (Qiagen Inc., Valencia, CA) en un cicladador térmico 9600 de Perkin-Elmer. Las condiciones de ciclación fueron 95°C durante 15 minutos para activar la Taq ADN polimerasa, seguido de 35 ciclos de 94°C durante 30 s., 64°C durante 30 s., 72°C durante 1 min. y una etapa de prolongación final a 72°C durante 10 min. Cantidades iguales de productos de PCR se electroforizaron en geles de agarosa al 2% y se observaron mediante tinte con bromuro de etidio. Todos los cebadores para RT-PCR abarcaban por lo menos 2 exones para impedir la contaminación por el ADN genómico. Para verificar la identidad de los productos de PCR, se clonaron en el vector pCR 2.1-TOPO (Invitrogen, Carlsbad, CA, USA) según las instrucciones del fabricante. Se secuenciaron las inserciones en ambas direcciones utilizando los cebadores específicos para el vector, con un secuenciador automático de ADN.

Expresión del tejido

El ARN total aislado de 26 tejidos humanos diferentes se adquirió en Clontech, Palo Alto, CA. Se preparó ADNc como se describió anteriormente para los experimentos en cultivo de tejido y se utilizó para las reacciones por PCR. Se ampliaron los ADNc del tejido a varias diluciones utilizando dos cebadores específicos para el gen (KLK15-F2 - SEC. ID n°: 49 y KLK15-R1 - SEC. ID n°: 48) (Tabla 1). Debido al alto grado de homología entre las calicreínas y a excluir la ampliación no específica, los productos de PCR se clonaron y se secuenciaron.

Tejidos de cáncer de próstata

Se extrajeron muestras de tejido prostático de 29 pacientes que se habían sometido a prostatectomía retropública radical para el adenocarcinoma prostático en el Charite University Hospital, Berlín, Alemania. Los pacientes no recibieron ninguna terapia hormonal antes de la cirugía. La utilización de estos tejidos con fines de investigación fue aprobada por el Ethics Committee of the Charite Hospital. Se extrajeron muestras de tejido de próstata reciente de las partes cancerosas y no cancerosas de las mismas próstatas que se habían extraído. Algunas piezas de tejido se disecaron inmediatamente después de la eliminación de la próstata y se almacenaron en nitrógeno líquido hasta el análisis. El análisis histológico de todas las piezas de tejido se realizó como se describió anteriormente (31), para asegurar que el tejido era maligno o benigno. Se pulverizaron los tejidos con un martillo bajo nitrógeno líquido y se extrajo el ARN como se describió anteriormente utilizando reactivo Trizol.

Análisis estadístico

Se realizó el análisis estadístico con el programa informático SAS (SAS Institute, Cary, NC). El análisis de las diferencias entre la expresión de KLK15 en tejidos no cancerosos frente a cancerosos del mismo paciente se realizó con el ensayo no paramétrico de McNemar. Se utilizó la distribución binómica para informatizar el nivel de significación. Los niveles del ARNm de KLK15 del tumor prostático se clasificaron cualitativamente en dos categorías (grupos bajo en KLK15 y alto en KLK15) y se analizaron las asociaciones entre el estado de KLK15 y otras variables utilizando la prueba exacta de Fisher.

Análisis estructural

Se realizó la alineación múltiple utilizando el paquete informático "Clustal X" y el programa de alineación múltiple disponible en el Baylor College of Medicine, Houston, TX, USA. Se realizaron estudios filogenéticos utilizando el paquete informático "Phylip". Se realizó el análisis de la matriz de distancia utilizando el programa "Neighbor-Joining/UPGMA" y se realizó el análisis de parsimonia utilizando el programa "Protpars". Se realizó el estudio de hidrofobia utilizando la lanzadera de búsqueda del Baylor College of Medicine. Se predijo el péptido señal utilizando el servidor "SignalP". Se realizó el análisis estructural de la proteína mediante el programa "SAPS" (análisis estructural de la secuencia proteica).

Resultados

Clonación del gen *KLK15*

Se creó previamente un mapa contiguo para el locus del gen de calicreína humana que se extiende desde el gen *KLK1* (centrómero) al gen *KLK14* (telómero) (7, 8, 11, 12, 27). A fin de investigar la presencia de otros genes de tipo calicreína centroméricos para *KLK1*, se obtuvo un clon de BAC (BC 781134) como se describe en materiales y métodos. Según la secuencia genómica publicada del antígeno específico de la próstata (PSA), y los genes de calicreína renal humana (*KLK1*), se diseñaron cebadores específicos para el gen para cada uno de estos genes (Tabla 1, SEC. ID n°: 47 a 50) y se desarrollaron protocolos de ampliación basados en la reacción en cadena de polimerasa (PCR) lo que permitió la generación de productos de PCR específico con ADN genómico como plantilla. La identificación por PCR del clon BAC por estos cebadores específicos para el gen indicó que este clon es positivo para *KLK1* pero negativo para PSA, confirmando, por lo tanto, qué suposición era centromérica con respecto a PSA.

Se previó una nueva serina proteasa supuesta a partir de la secuencia de este clon mediante programas informáticos como se describió anteriormente (12). Se digirió este clon, se transfirió a una membrana y se hibridó con cebadores específicos para el gen *KLK15* supuesto (según la secuencia prevista) y los fragmentos positivos se subclonaron y secuenciaron para verificar la estructura del gen supuesto. La secuencia supuesta del gen se arremetió a continuación contra la base de datos EST humana y se identificaron dos clones EST (n° de registro del GenBank AW274270 y n° AW205420). Estos dos clones eran 99% idénticos al último exón y la zona 3' no traducida del gen y a los segundos extremos EST con una prolongación de 17 nucleótidos de adenina (A) que no se encontraban en la secuencia genómica, verificando de este modo el extremo 3' del gen y la posición de la cola poli A.

Para identificar la estructura del ARNm completo del gen y para determinar los límites exón/intrón, se realizaron reacciones de PCR utilizando cebadores situados en diferentes exones predichos por ordenador, utilizando un panel de 26 ADNc de tejido humano como plantilla. Se secuenciaron los productos de PCR. Dos de estos cebadores (*KLK15-F1* - SEC. ID n°: 47 y *KLK15-R1* - SEC. ID n°: 48)) (Tabla 1) fueron capaces de ampliar la zona de codificación completa del gen de diferentes tejidos. La comparación del ARNm con la estructura genómica indicó la presencia de un gen formado por cinco exones de codificación con 4 intrones de intervención. La traducción de la secuencia del ARNm en todos los posibles marcos de lectura puso de manifiesto la presencia de sólo un marco que da una cadena polipeptídica ininterrumpida, que también contiene los motivos estructurales muy conservados de las calicreínas, como se expone a continuación.

Caracterización estructural del gen *KLK15*

Como se muestra en la Figura 1, el gen *KLK15* está formado por 5 exones de codificación y 4 intrones de intervención aunque, como con otros genes de la calicreína, la presencia de exón(es) adicional(es) no traducido(s) corriente arriba podría no estar regulada (17, 32, 33). Todas las secuencias de corte y empalme de exón/intrón conforman la secuencia de consenso para las secuencias de corte y empalme eucarióticas (34). El gen sigue además estrictamente las propiedades estructurales comunes de otros miembros de la familia multigén de calicreína humana, descrita más adelante. La zona de codificación de la proteína prevista del gen está formada por 771 bp, que codifican un polipéptido de 256 aminoácidos deducido con un peso molecular previsto de 28,1 kDa. El codón de iniciación de la traducción potencial es compatible con la secuencia de consenso Kozak (35), además, existe una purina en la posición (-3) que se produce en el 97% de los ARNm de vertebrado (36). Debe también observarse, que como la mayoría de otros genes de tipo calicreína, *KLK15* no tiene el nucleótido G de consenso en la posición (+4).

Los nucleótidos 7764 a 7769 (ATTAAA) (SEC. ID n°: 40) se parecen estrechamente a una señal de consenso de poliadenilación (37) si son seguidos, después de 17 nucleótidos por la cola poli A. Ninguna otra señal de poliadenilación es discernible en la zona 3' no traducida, lo que sugiere que la secuencia anterior es la señal de poliadenilación existente. Aunque AATAAA (SEC. ID n°: 40), es muy conservada, se producen variantes naturales, y la secuencia AATAAA (SEC. ID n°: 40) se describe que tiene lugar como variante natural de poliadenilación en el 12% de las secuencias del ARNm del vertebrado (38). La presencia de ácido glutámico (E) en la posición 203 sugiere que *KLK15* será probablemente una única especificidad de sustrato. PSA tiene un resto de serina (S) en la posición correspondiente y presenta actividad de tipo quimotriptica. Muchas otras calicreínas normalmente tienen aspartato (D) en esta posición lo que indica una actividad similar a la tripsina (Figura 2) (6).

Aunque la secuencia de la proteína de *KLK15* es única, el análisis comparativo puso de manifiesto que presenta un grado considerable de homología con otros miembros de la familia del multigén calicreína. *KLK15* presenta el 51% de identidad de la proteína y el 66% de similitud con la serina proteasa similar a la tripsina (TLSP) y el 49%, 48% de identidad con la neuropsina y las proteínas *KLK-L3*, respectivamente. El análisis de hidrofobia puso de manifiesto que la zona amino-terminal es completamente hidrófoba (Figura 3), acorde con la posibilidad de que esta zona pueda albergar una secuencia fenol, análoga a otras serina proteasas. El análisis informático de la secuencia de la proteína de *KLK15* previó una secuencia de escisión entre los aminoácidos 16 y 17 (TAA-QD). La alineación de la secuencia (Figura 2) puso de manifiesto también otra secuencia de escisión potencial (Lys²¹), en un sitio homólogo a la secuencia de activación de otras serina proteasas [lisina (K) o arginina (R) está presente en la mayoría de los casos] (39). Varias zonas hidrófobas uniformemente distribuidas en todo el polipéptido *KLK15* son corrientes con una proteína globular, similar a otras calicreínas y serina proteasas. Por lo tanto, como es el caso con otras calicreínas, *KLK15* se traslada

supuestamente como un precursor inactivo de las preproenzima de 256 aminoácidos. Prepro-KLK15 tiene 21 restos adicionales que constituyen la pre-zona (péptido señal formado por 16 restos) y el propéptido (5 restos).

La zona punteada en la Figura 2 indica un bucle de 11 aminoácidos característico de las calicreínas clásicas (PSA, KLK1 y KLK2) pero no se encuentra en KLK15 y otros miembros de la familia del multigén con calicreína (10, 11, 13, 14). Sin embargo, KLK15 tiene un único bucle de 8 aminoácidos (HNEPGTAG) (SEC. ID nº: 10) en las posiciones 148 a 155, no hallado en ninguna otra calicreína (Figura 2). Se han descrito veintinueve aminoácidos “invariantes” que rodean la secuencia activa de las serina proteasas (40). De éstos, veintiocho se conservan en KLK15. Uno de los aminoácidos no conservados (Ser¹⁷³ en lugar de Pro) se encuentra también en la prostasa, proteínas KLK-L2 y KLK-L5, y representa un cambio evolutivo conservado en una proteína del mismo grupo, según los estudios de evolución de la proteína (41). Doce restos de cisteína se presentan en la supuesta proteína de KLK15 madura; diez de ellos se conservan en todas las calicreínas y sería de esperar que formen puentes disulfuro. Los otros dos (C131 y C243) no se encuentran en PSA, KLK1, KLK2 o KLK-L4, sin embargo, se encuentran en posiciones similares en todos los demás genes de calicreína y es de esperar que formen un enlace disulfuro adicional.

Para predecir la relevancia filogenética de la proteína de KLK15 con otras serina proteasas, se alinearon las secuencias de aminoácidos conjuntamente utilizando el método del grupo de pares ponderados con la media aritmética (UPGMA) y los métodos Neighbor-Joining de la matriz de distancia, y el método de parsimonia “Protpars”. Todos los árboles filogenéticos obtenidos coincidieron en que otras serinas proteasas (no calicreínas) pueden agruparse como un grupo independiente, indicando que las calicreínas representan una etapa independiente en la evolución de las serinas proteasas. KLK15 se agrupó con la KLK-L3 y TLSP (Figura 4) y las calicreínas clásicas (hK1, hK2 y PSA) se agruparon en todos los árboles, lo que sugiere que la separación entre las calicreínas clásicas y los genes de tipo calicreína aparecieron al principio de la evolución, acorde con las sugerencias de los estudios anteriores (13).

Variantes de corte y empalme del gen KLK15

La identificación por PCR de los transcritos de KLK15 que utilizan cebadores específicos del gen (KLK15-F2 - SEC. ID nº: 49 y KLK15-R2 - SEC. ID nº: 50) (Tabla 1) puso de manifiesto la presencia de 3 bandas en la mayoría de los ADNc del tejido examinados (Figura 6). Estas bandas se purificaron en gel, se clonaron y se secuenciaron. La banda superior representa la forma clásica del gen y la banda inferior es la variante 3 de corte y empalme (Figura 7). La banda media representa otras dos variantes de corte y empalme. La digestión de la restricción del producto de la PCR de la banda media con Stu I, seguida de la separación en gel, purificación y secuenciado puso de manifiesto que está compuesta de las variantes 1 y 2 de corte y empalme que tienen aproximadamente la misma longitud (variante 1 de corte y empalme tiene el exón 4 (137 bp) pero está perdiendo 118 bp del exón 3, mientras que la variante de corte y empalme 2 tiene otros 118 bp del exón 3 pero está perdiendo el exón 4. Todas las variantes de corte y empalme es de esperar que codifiquen productos de proteína truncada (Figura 5).

Localización de cromosoma del gen KLK15

El estudio del análisis de restricción de numerosos clones BAC solapantes que abarcan el locus de calicreína humana seguido por comparación con la cartografía de restricción de EcoR1 del área (disponible en la secuencia de la red LLNL) permitió la identificación de un clon BAC (BC 25479) que es adyacente teloméricamente a BC 781134 (que alberga el gen KLK15). La explosión de las secuencias de los dos clones demostró que los extremos de estos clones son solapantes. Identificando la posición de los genes KLK1, KLK3 y KLK15 a lo largo de estos clones, la posición relativa y la dirección de la transcripción de estos tres genes se definió exactamente. KLK1 es el más centromérico y su dirección de transcripción es desde el telómero hasta el centrómero, seguido por KLK15, que es más telomérico y transcribe en la misma dirección. La distancia entre los dos genes es 1501 bp de longitud. El gen KLK3 es más telomérico, situado a una distancia de 23.335 del KLK15, y se transcribe en la dirección opuesta (Figura 6). Estos resultados son acordes con los informes previos donde la distancia entre KLK3 y KLK1 se estimó que era aproximadamente ~ 31 Kb (6,27).

Expresión del tejido y regulación hormonal del gen KLK15

Como se muestra en la Figura 7, el gen KLK15 se expresa a niveles mayores en la glándula tiroides. Los niveles menores de expresión se observan también en la próstata, glándulas salivales y suprarrenales, colon, testículos y riñón. A fin de verificar la especificidad de RT-PCR, se clonaron productos de PCR representativos y se secuenciaron. La Figura 8 demuestra que el gen KLK15 está regulado por incremento por hormonas esteroideas en la línea celular humana LNCaP del cáncer de próstata.

Expresión de KLK15 en el cáncer de próstata

La expresión del gen KLK15 en tejidos prostáticos normales y cancerosos se examinó por RT-PCR. Se incluyó actina como gen de referencia para asegurar la calidad y cantidad del ADNc utilizado. A fin de examinar la expresión relativa del gen KLK15 en tejidos normales comparados con malignos, se examinaron 29 pares de tejidos prostáticos. Cada par representaba el tejido normal y canceroso extraído del mismo paciente. Los resultados se resumen en la Tabla 2. Trece de cada 29 pacientes presentaban expresión de KLK15 significativamente mayor en el tejido canceroso y solamente tres presentaban la expresión de KLK15 mayor en los tejidos no cancerosos que en los cancerosos. El análisis mediante la prueba de McNemar indicaba que las diferencias entre los tejidos normales y cancerosos son

estadísticamente significativas ($P=0,021$). Debido al pequeño número de casos, se utilizó la distribución binómica para informatizar el nivel de significación. Los pacientes de cáncer de próstata se clasificaron además en dos grupos: (a) positivos a la expresión de KLK15 ($N=21$) y (b) negativos a la expresión de KLK15 (o muy bajos) ($N=8$). Cuando la asociación de la expresión de KLK15 se comparaba con las variables del pronóstico clínicopatológico se halló la mayor expresión de KLK15 que era más frecuente en pacientes con enfermedad en la etapa terminal y tumores de mayor grado (Tabla 3).

Exposición

Las calicreínas son un subgrupo de serina proteasas. El término “calicreína” se utiliza normalmente para describir una enzima que actúa en una molécula precursora (cininógeno) para la liberación de un péptido bioactivo (cinina) (3, 42). Sin embargo, la expresión genérica “calicreína tisular” no se limita a la definición operativa de la enzima. Esta expresión se utiliza actualmente para describir un grupo de enzimas con estructura génica y proteica muy conservada que también se localiza conjuntamente en el mismo locus cromosómico. Entre los tres genes de la calicreína humana clásicos, solamente el KLK1 codifica una proteína con potente actividad de cininogenasa. Las enzimas codificadas por los genes KLK2 y KLK3 presentan actividad de cininogenasa muy débil. Los 14 miembros ya clonados de la familia del gen humano de la calicreína tienen numerosas similitudes (7, 11) como se muestra a continuación:

- Todos los genes localizan la misma zona cromosómica (19q13.3-q13.4)
- Todos los genes codifican supuestas serina proteasas con una triada catalítica conservada en las posiciones apropiadas, es decir, histidina cerca del extremo del segundo exón de codificación, ácido aspártico en el medio del tercer exón y serina al comienzo del quinto (último exón).
- Todos los genes tienen cinco exones de codificación (algunos miembros contienen uno o más de 5 exones no traducidos).
- Los tamaños del exón de codificación son similares o idénticos.
- Las fases del intrón están completamente conservadas.
- Todos los genes tienen homologías de secuencia significativas en los niveles del ADN y de aminoácidos (30 al 80%).
- Muchos de estos genes están regulados por hormonas esteroideas.

Las Figuras 2 y 8 demuestran que los genes KLK15 recién identificados comparten todas las similitudes anteriores y es por lo tanto un nuevo miembro de la familia del multigén humano de calicreína. Este gen se denominó KLK15.

Muchos genes de calicreína están relacionados con la patogénesis de las enfermedades humanas, dependiendo en el tejido de su expresión primaria. El gen KLK1 está involucrado en muchos procesos patológicos, incluyendo la inflamación (3), hipertensión (44), nefritis renal y renopatía diabética (45, 46). Las conexiones de HSCCE (KLK7) con las enfermedades cutáneas, incluyendo la queratinización patológica y la psoriasis, se han descrito ya (47, 48). Little *et al.* sugirieron que la zima (KLK6) puede ser amiloidogénica y puede desempeñar una función en desarrollo de la enfermedad de Alzheimer (14). Existen otros informes que describen la relación de la expresión de la neuropsina (KLK8) con las enfermedades del sistema nervioso central, incluyendo la epilepsia (49, 50). Al expresarse principalmente en el tiroides, KLK15 puede desempeñar una función importante en la fisiología y patofisiología normales de esta glándula. Entre todas las demás calicreínas descubiertas, muchas se expresan en el tiroides pero ninguna a los niveles más elevados en este tejido (7, 11).

El gen KLK15 está regulado por incremento, en el nivel del ARNm, en un subconjunto de cánceres de próstata. Las distribuciones del estado de la expresión cualitativa de KLK15 (alto o bajo) entre subgrupos de pacientes que se diferencian en la etapa de la enfermedad, grado del tumor y puntuación de Gleason indicaron que la expresión elevada de KLK15 se encontraba más frecuentemente en los tumores de grado 3 así como en la etapa III y puntuación de Gleason >6 pacientes. Estos descubrimientos indican que la sobreexpresión de KLK15 está relacionada con las formas más agresivas de la enfermedad y puede ser un indicador de pronóstico escaso (Tabla 3).

Existen actualmente pruebas crecientes de que muchas calicreínas y genes de tipo calicreína están relacionados con el cáncer. PSA es el mejor marcador para el cáncer de próstata hasta ahora (20). Los informes recientes sugieren que hK2 (codificado por el gen KLK2) puede ser otro marcador de diagnóstico útil para el cáncer de próstata (21, 51). NES1 (KLK10) parece ser un nuevo gen supresor tumoral (23). El gen de la zima (KLK6) se demostró que se expresaba de forma diferenciada en los tumores de mama y de ovario primarios (24) y la enzima quimocriptica de la capa córnea humana (HSCCE, KLK7) se ha demostrado que se expresa a niveles anormalmente elevados en el cáncer de ovario (25). Otro gen de tipo calicreína identificado recientemente, denominado provisionalmente el gen 14 expresado de forma diferenciada asociado al tumor (TADG-14)/neuropsina (KLK8) se descubrió que se sobreexpresaba aproximadamente en el 60% de los tejidos de cáncer de ovario (26). Pprostasa/KLK-L1/ (KLK4), otro gen de tipo

calicreína descubierto recientemente, se especula que está relacionado con el cáncer de próstata (13). Dos calicreínas recién descubiertas, KLK-L4 (KLK13) y KLK-L5 (KLK12), se ha descubierto también que están reguladas por disminución en el cáncer de mama (10). Por lo tanto, la nueva extensa bibliografía sugiere múltiples relaciones de varios genes de calicreína con muchas formas de cáncer humano.

La existencia de múltiples formas de ARNm cortadas y empalmadas alternativamente es frecuente entre las calicreínas. Distintas especies de ARN se transcriben en el gen PSA, además del transcrito principal de 1,6 kb (19, 52, 53). Asimismo, Reigman *et al.* describieron la identificación de dos formas cortadas y empalmadas alternativamente del gen de calicreína 2 (KLK2) glandular humana (54). Un nuevo transcrito del gen de la calicreína tisular (KLK1) se aisló también del colon (55). La neuropsina, un gen de tipo calicreína identificado recientemente, se descubrió que presenta dos formas cortadas y empalmadas alternativamente, además de la forma principal (26, 56). Se descubrió también que KLK-L4 presenta diferentes formas cortadas y empalmadas alternativamente (10). Debido a que las variantes de corte y empalme de KLK15 tienen una secuencia 5' idéntica requerida para la traducción, secreción y activación, es posible suponer que codifican una proteína segregada (53).

En conclusión, un nuevo elemento de la familia de genes de la calicreína humana, KLK15, se ha caracterizado porque cartografía el locus de calicreína humana (cromosoma 19q13.3-q13.4). Este gen presenta tres formas de corte y empalme relacionadas además de la forma clásica. KLK15 se expresa en una variedad de tejidos pero principalmente en el tiroides, parece estar regulada por incremento en las formas más agresivas del cáncer de próstata y su expresión está influenciada por las hormonas esteroides. Dado que unas pocas calicreínas se utilizan ya como marcadores tumorales valiosos, KLK15 puede también encontrar aplicaciones clínicas similares.

TABLA 1

Cebadores utilizados para la ampliación por PCR genómica

Gen	Denominación y secuencia del cebador		Nº de registro en GenBank
KLK1	KLK1-A:	ATC CCT CCA TTC CCA TCT TT	L10038
	KLK1-B:	CAC ATA CAA TTC TCT GGT TC	
KLK2	KLK2-A:	AGT GAC ACT GTC TCA GAA TT	M18157
	KLK2-B:	CCC CAA TCT CAC GAG TGC AC	
PSA	E5-A:	GTC GGC TCT GGA GAC ATT TC	M27274
	E5-B:	AAC TGG GGA GGC TTG AGT C	
KLK15	KLK15-F1	CTC CTT CCT GCT GGC ATC CA	AF242195
	KLK15-R1	ATC ACA CGG GTG GTC ATG TG	
	KLK15-F2	CAA GTG GCT CTC TAC GAG CG	
	KLK15-R2	GAC ACC AGG CTT GGT GGT GT	

*Todos los cebadores se presentan en la dirección 5' → 3'.

TABLA 2

Expresión de KLK15 en 29 pares de tejidos prostáticos cancerosos y no cancerosos

Expresión de KLK15	Número de pacientes	Valor P*
Más alta en el cáncer frente a normal	13	
Más baja en el cáncer frente a normal	3	
Expresión alta pero aprox. igual en ambos tejidos	8	
Expresión baja (o no) pero aprox. igual en ambos tejidos	5	0,021

*El valor P se calculó por la prueba de McNemar utilizando la distribución binómica.

TABLA 3

Relación entre la expresión de KLK15 y otras variables clínicopatológicas en 29 pacientes con cáncer de próstata primario

Variable	Pacientes	Número de pacientes (%)		Valor P*
		Negativos a KLK15	Positivos a KLK15	
Etapas				
I/II	20	8 (40)	12 (60)	0,033
III	9	0 (0)	9 (100)	
Grado				
G1/2	23	8 (34,8)	15 (65,2)	0,15
G3	6	0 (0)	6 (100)	
Puntuación de Gleason				
≤6	22	7 (31,8)	15 (62,2)	0,14
>6	6	0 (0)	6 (100)	
desconocida	1			

*Prueba exacta de Fisher.

Citas completas de las referencias citadas en memoria

1. Kraut, H., K., F. E., y Werle, E. (1930) *Physiol. Chem.* 189,97-106
2. Werle, E. (1934) *Biochem. Z* 269, 415-434
3. Clements, J. (1997) in *The Kinin System* (Farmer, S., ed), pp. 71-97, *Academic Press*, New York
4. Evans, B. A., Drinkwater, C. C., y Richards, R. 1. (1987) *J. Biol. Chem.* 262, 8027-34
5. Ashley, P. L., y MacDonald, R. J. (1985) *Biochemistry* 24, 4520-7
6. Riegman, P. H., Vlietstra, R. J., Suurmeijer, L., Cleutjens, C. B., y Trapman, J. (1992) *Genomics* 14,6-11
7. Diamandis, E. P., Yousef, G. M., Luo, L. Y., Magklara, A., y Obiezu, C. V. (2000) *Trends Endocrinol. Metab.* 11, 54-60
8. Yousef, G. M., Obiezu, C. V., Luo, L. Y., Black, M. H., y Diamandis, E. P. (1999) *Cancer Res.* 59, 4252-6
9. Yousef, G. M., y Diamandis, E. P. (1999) *J. Biol. Chem.* 274, 37511-6
10. Yousef, G. M., Chang, A., y Diamandis, E. P. (2000) *J Biol Chem* 275, 11891-8.
11. Yousef, G. M., y Diamandis, E. P. (2000) *Genomics* 65, 184-194
12. Yousef, G. M., Luo, L. Y., y Diamandis, E.P. (1999) *Anticancer Res.* 19, 2843-52
13. Nelson, P. S., Gan, L., Ferguson, C., Moss, P., Gelinas, R., Hood, L., y Wang, K. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96,3114-9
14. Little, S. P., Dixon, E. P., Norris, F., Buckley, W., Becker, G. W., Johnson, M., Dobbins, J. R., Wyrick, T., Miller, J. R., MacKellar, W., Hepburn, D., Corvalan, J., McClure, D., Liu, X., Stephenson, D., Clemens, J., y Johnstone, E. M. (1997) *J. Biol. Chem.* 272, 25135-42
15. Liu, X. L., Wazer, D. E., Watanabe, K., y Band, V. (1996) *Cancer Res.* 56, 3371-9
16. Hansson, L., Stromqvist, M., Backman, A., Wallbrandt, P., Carlstein, A., y Egelrud, T. (1994) *J. Biol. Chem.* 269, 19420-6
17. Yoshida, S., Taniguchi, M., Hirata, A., y Shiosaka, S. (1998) *Gene* 213(1-2), 9-16
18. Stephenson, S. A., Verity, K., Ashworth, L. K., y Clements, J. A. (1999) *J. Biol. Chem.* 274,23210-4

ES 2 266 245 T3

19. **Riegman, P. H., Vlietstra, R. J., van der Korput, J. A., Romijn, J. C., y Trapman, J.** (1989) *Biochem. Biophys. Res. Commun.* 159, 95-102
20. **Diamandis, E. P.** (1998) *Trends Endocrinol. Metab.* 9, 310-316
21. **Stenman, U. H.** (1999) *Clin. Chem.* 45, 753-4.
22. **Partin, A. W., Catalona, W. J., Finlay, J. A., Darte, C., Tindall, D. J., Young, C. Y., Klee, G. G., Chan, D. W., Rittenhouse, H. G., Wolfert, R. L., y Woodrum, D. L.** (1999) *Urology* 54, 839-45
23. **Goyal, J., Smith, K. M., Cowan, J. M., Wazer, D. E., Lee, S. W., y Band, V.** (1998) *Cancer Res.* 58, 4782-6
24. **Anisowicz, A., Sotiropoulou, G., Stenman, G., Mok, S. C., y Sager, R.** (1996) *Mol. Med.* 2, 624-36
25. **Tanimoto, H., Underwood, L. J., Shigemasa, K., Yan Yan, M. S., Clarke, J., Parmley, T. H., y O'Brien, T. J.** (1999) *Cancer* 86, 2074-82
26. **Underwood, L. J., Tanimoto, H., Wang, Y., Shigemasa, K., Parmley, T. H., y O'Brien, T. J.** (1999) *Cancer Res.* 59(17), 4435-9
27. **Yousef, G. M., Chang, A., y Diamandis, E. P.** (2000) submitted
28. **Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., y Lipman, D. J.** (1997) *Nucleic Acids Res.* 25(17), 3389-402
29. **Lennon, G., Auffray, C., Polymeropoulos, M., y Soares, M. B.** (1996) *Genomics* 33, 151-2
30. **Sambrook, J., Fritsch, E. F., y Maniatis, T.** (1989) *Molecular Cloning: A Laboratory Manual*, Second Edition Ed.. Cold Spring Harbor Laboratory, NY
31. **Meyer, A., Jung, K., Lein, M., Rudolph, B., Schnorr, D., y Loening, S. A.** (1997) *Int. J. Cancer* 74, 630-6
32. **Luo, L., Herbrick, J. A., Scherer, S. W., Beatty, B., Squire, J., y Diamandis, E. P.** (1998) *Biochem. Biophys. Res. Commun.* 247(3), 580-6
33. **Yousef, G. M., Luo, L. Y., Scherer, S. W., Sotiropoulou, G., y Diamandis, E. P.** (1999) *Genomics* 62(2), 251-9
34. **Iida, Y.** (1990) *J. Theor. Biol.* 145(4), 523-33
35. **Kozak, M.** (1991) *J. Cell Biol.* 115(4), 887-903
36. **Kozak, M.** (1987) *Nucleic Acids Res.* 15(20), 8125-48
37. **Proudfoot, N. J., y Brownlee, G. G.** (1976) *Nature* 263, 211-4
38. **Sheets, M. D., Ogg, S. C., y Wickens, M. P.** (1990) *Nucleic Acids Res.* 18(19), 5799-805
39. **Keil, B.** (1971) in *The enzymes* (P.D. Boyer, E., ed) Vol. 3, 3rd Ed., pp. 249-275, *Academic Press*, New York
40. **Dayhoff, M. O.** (1978) *Natl. Biomed. Res. Found.* 5, 79-81
41. **Miyata, T., Miyazawa, S., y Yasunaga, T.** (1979) *J. Mol. Evol.* 12(3), 219-36
42. **Bhoola, K. D., Figueroa, C. D., y Worthy, K.** (1992) *Pharmacol. Rev.* 44(1), 1-80
43. **Diamandis, E. P., Yousef, G., Clements, J. et al.,** (2000) *Clin. Chem.* In press
44. **Margolius, H. S., Horwitz, D., Pisano, J. J., y Keiser, H. R.** (1974) *Circ. Res.* 35(6), 820-5
45. **Jaffa, A. A., Chai, K. X., Chao, J., Chao, L., y Mayfield, R. K.** (1992) *Kidney Int.* 41(4), 789-95
46. **Cumming, A. D., Walsh, T., Wojtacha, D., Fleming, S., Thomson, D., y Jenkins, D. A.** (1994) *Clin. Sci.* 87 (1), 5-11
47. **Sondell, B., Dyberg, P., Anneroth, G. K., Ostman, P. O., y Egelrud, T.** (1996) *Acta. Derm. Venereol.* 76(3), 177-81

48. Ekholm, E., y Egelrud, T. (1999) *Arch. Dermatol. Res.* 291(4), 195-200

49. Momota, Y., Yoshida, S., Ito, J., Shibata, M., Kato, K., Sakurai, K., Matsumoto, K., y Shiosaka, S. (1998) *Eur. J. Neurosci.* 10(2), 760-4

50. Kishi, T., Kato, M., Shimizu, T., Kato, K., Matsumoto, K., Yoshida, S., Shiosaka, S., y Hakoshima, T. (1999) *J. Biol. Chem.* 274, 4220-4

51. Black, M. H., Magldara, A., Obiezu, C. V., Melegos, D. N., y Diamandis, E. P. (1999) *Clin. Chem.* 45(6 Pt 1), 790-9

52. Riegman, P. H., Klaassen, P., van der Korput, J. A., Romijn, J. C., y Trapman, J. (1988) *Biochem. Biophys. Res. Commun.* 155(1), 181-8

53. Heuze, N., Olayat, S., Gutman, N., Zani, M. L., y Courty, Y. (1999) *Cancer Res.* 59, 2820-4

54. Riegman, P. H., Vlietstra, R. J., van der Korput, H. A., Romijn, J. C., y Trapman, J. (1991) *Mol. Cell Endocrinol.* 76(1-3), 181-90

55. Chen, L. M., Murray, S. R., Chai, K. X., Chao, L., y Chao, J. (1994) *Braz. J. Med. Biol. Res.* 27(8), 1829-38

56. Mitsui, S., Tsuruoka, N., Yamashiro, K., Nakazato, H., y Yamaguchi, N. (1999) *Eur. J. Biochem.* 260(3), 627-34.

REIVINDICACIONES

1. Proteína aislada que comprende una secuencia de aminoácidos de la SEC. ID. nº: 6, 7, 8 ó 9.

2. Molécula de ácido nucleico aislada que codifica una proteína según la reivindicación 1, o una secuencia de ácidos nucleicos complementaria a ésta, o una molécula de ácido nucleico que presenta por lo menos 80%, 85%, 90%, 95% o 99% de identidad con ésta.

3. Molécula de ácido nucleico aislada según la reivindicación 2, que presenta una secuencia de ácidos nucleicos de la SEC. ID nº: 1, 2, 3, 4 ó 5 en la que T puede ser además U, o una secuencia de ácidos nucleicos complementaria a ésta, o una molécula de ácido nucleico que presenta por lo menos 80%, 85%, 90%, 95% o 99% de identidad con ésta.

4. Vector que comprende una molécula de ácido nucleico según la reivindicación 2 ó 3.

5. Célula huésped que comprende una molécula de ácido nucleico según la reivindicación 2 ó 3.

6. Método para la preparación de una proteína según la reivindicación 1, que comprende las etapas siguientes:

- (a) transferir un vector de expresión según la reivindicación 4 a una célula huésped;
- (b) seleccionar las células huésped transformadas de entre las células huésped no transformadas;
- (c) cultivar una célula huésped transformada seleccionada en condiciones que permitan la expresión de la proteína; y
- (d) aislar la proteína.

7. Anticuerpo que presenta especificidad contra un epítipo de una proteína según la reivindicación 1.

8. Anticuerpo según la reivindicación 7, marcado con una sustancia detectable y utilizado para detectar el polipéptido en muestras biológicas, tejidos y células.

9. Sonda que comprende una secuencia que codifica una proteína según la reivindicación 1.

10. Kit de ensayo para el diagnóstico de una enfermedad asociado a una proteína según la reivindicación 1, mediante la determinación de la presencia de una molécula de ácido nucleico según cualquiera de las reivindicaciones anteriores o una proteína según cualquiera de las reivindicaciones anteriores.

11. Kit de ensayo según la reivindicación 10, en el que la enfermedad es el cáncer.

12. Kit de ensayo según la reivindicación 11, en el que el cáncer es el cáncer de próstata.

13. Método para identificar una sustancia que se asocia con una proteína según la reivindicación 1, que comprende las etapas siguientes:

- (a) hacer reaccionar la proteína con al menos una sustancia que puede asociarse potencialmente con la proteína, en condiciones que permiten la asociación entre la sustancia y la proteína, y
- (b) eliminar o detectar la proteína asociada con la sustancia, en el que la detección de la proteína y la sustancia asociadas indica que la sustancia se asocia con la proteína.

14. Método para evaluar un compuesto por su capacidad para modular la actividad biológica de una proteína según la reivindicación 1, que comprende proporcionar la proteína con una sustancia que se asocia con la proteína y un compuesto de ensayo en condiciones que permitan la formación de complejos entre la sustancia y la proteína, y eliminar y/o detectar los complejos.

15. Método para identificar los inhibidores de la interacción de una proteína según la reivindicación 1 y una sustancia que interactúa con la proteína que comprende las etapas siguientes:

- (a) proporcionar una mezcla de reacción que comprende la proteína y una sustancia que se une a la proteína, o por lo menos una parte de cada una que interactúa;
- (b) poner en contacto la mezcla de reacción con uno o más compuestos de ensayo;
- (c) identificar los compuestos que inhiben la interacción de la proteína y la sustancia.

16. Método para la detección de una molécula de ácido nucleico que codifica una proteína según la reivindicación 1 en una muestra biológica que comprende las etapas siguientes:

- (a) hibridar una molécula de ácido nucleico según la reivindicación 2 ó 3 con ácidos nucleicos de la muestra biológica, formando de esta manera un complejo de hibridación; y
- (b) detectar el complejo de hibridación en el que la presencia del complejo de hibridación está correlacionada con la presencia de una molécula de ácido nucleico que codifica la proteína en la muestra biológica.

17. Método según la reivindicación 16, en el que los ácidos nucleicos de la muestra biológica se amplían mediante la reacción en cadena de la polimerasa antes de la etapa de hibridación.

18. Método *in vivo* para controlar la evolución del cáncer de próstata en un individuo, que comprende las etapas siguientes:

- (a) poner en contacto una cantidad de un anticuerpo que se une a una proteína, según la reivindicación 1, con una muestra del individuo de manera que formen un complejo binario que comprende el anticuerpo y la proteína en la muestra;
- (b) determinar o detectar la presencia o cantidad de formación de complejo en la muestra;
- (c) repetir las etapas (a) y (b) en un punto posterior en el tiempo; y
- (d) comparar el resultado de la etapa (b) con el resultado de la etapa (c), en el que una diferencia en la cantidad de formación de complejo es indicativa de la evolución del cáncer en dicho individuo.

19. Utilización de una proteína según la reivindicación 1, para la preparación de un medicamento destinado al tratamiento del cáncer de próstata.

20. Composición farmacéutica que comprende una o más moléculas de ácido nucleico o proteínas según cualquiera de las reivindicaciones anteriores, y un portador, excipiente o diluyente farmacéuticamente aceptable.

21. Utilización de una o más moléculas de ácido nucleico o proteínas según cualquiera de las reivindicaciones anteriores, para la preparación de una composición farmacéutica destinada al tratamiento del cáncer de próstata.

22. Método para realizar el descubrimiento del fármaco que comprende:

- (a) proporcionar los agentes identificados por su capacidad para inhibir o potenciar la interacción de una proteína según la reivindicación 1 y una sustancia que se une a la proteína;
- (b) realizar el perfil terapéutico de los agentes identificados en la etapa (a) o análogos adicionales de los mismos, para la eficacia y la toxicidad en animales; y
- (c) determinar una formulación para una preparación farmacéutica que comprende uno o más agentes identificados en la etapa (b) que presentan un perfil terapéutico aceptable.

23. Vacuna para estimular o potenciar en un sujeto para el que la vacuna es la producción de anticuerpos administrada dirigida contra una proteína tal como en la reivindicación 1.

24. Utilización de una proteína según la reivindicación 1, para la preparación de una vacuna destinada a estimular o potenciar en un sujeto la producción de anticuerpos dirigida contra la proteína.

Figura 1

TGGATTCTCTCACTCCCTCCCCAGACTGCAGCCGAACCCTGGTCCCTCCTCCACA
 (ATG) TGG CTT CTC CTC ACT CTC TCC TTC CTG CTG GCA TCC ACA
 M W L L L T L S F L L A S T
 G gtgaggtggccccaggagggggccaggtctgtgggagcaggtg.....
 ...Intron1.....gcatcctctacccttctcttag CA GCC CAG
 A A Q
 GAT GGT GAC AAG TTG CTG GAA GGT GAC GAG TGT GCA CCC CAC
 D G D K L L E G D E C A P H
 TCC CAG CCA TGG CAA GTG GCT CTC TAC GAG CGT GGA CGC TTT
 S Q P W Q V A L Y E R G R F
 AAC TGT GGC GCT TCC CTC ATC TCC CCA CAC TGG GTG CTG TCT
 N C G A S L I S P H W V L S
 GCG GCC CAC TGC CAA AGC CG gtatgaaggcaggggctcagggctcctga
 A A [H] C Q S R
 gggg.....Intron 2cgcaactccactggcgggaaa
 accactcgccccgcacag C TTC ATG AGA GTG CGC CTG GGA GAG CAC
 F M R V R L G E H
 AAC CTG CGC AAG CGC GAT GGC CCA GAG CAA CTA CGG ACC ACG
 N L R K R D G P E Q L R T T
 TCT CGG GTC ATT CCA CAC CCG CGC TAC GAA GCG CGC AGC CAC
 S R V I P H P R Y E A R S H
 CGC AAC GAC ATC ATG TTG CTG CGC CTA GTC CAG CCC GCA CGC
 R N [D] I M L L R L V Q P A R
 CTG AAC CCC CAG GTG CGC CCC GCG GTG CTA CCC ACG CGT TGC
 L N P Q V R P A V L P T R C
 CCC CAC CCG GGG GAG GCC TGT GTG GTG TCT GGC TGG GGC CTG
 P H P G E A C V V S G W G L
 GTG TCC CAC AAC GAG CCT GGG ACC GCT GGG AGC CCC CGG TCA
 V S H N E P G T A G S P R S
 CAA G gtgcgtgaaaggatggagctggat.....Intron 3.....
 Q
 ctccaagtccactgtcttccccag TG AGT CTC CCA GAT ACG TTG CAT
 V S L P D T L H
 TGT GCC AAC ATC AGC ATT ATC TCG GAC ACA TCT TGT GAC AAG
 C A N I S I I S D T S C D K
 AGC TAC CCA GGG CGC CTG ACA AAC ACC ATG GTG TGT GCA GGC
 S Y P G R L T N T M V C A G
 GCG GAG GGC AGA GGC GCA GAA TCC TGT GAG gtcagagcctagagg
 A E G R G A E S C E
 ggccatcaggcggaagaagaggg.....Intron 4.....cct
 gagacccccctcttttccccacag GGT GAC TCT GGG GGA CCC CTG GTC
 G D [S] G G P L V
 TGT GGG GGC ATC CTG CAG GGC ATT GTG TCC TGG GGT GAC GTC
 C G G I L Q G I V S W G D V
 CCT TGT GAC AAC ACC ACC AAG CCT GGT GTC TAT ACC AAA GTC
 P C D N T T K P G V Y T K V
 TGC CAC TAC TTG GAG TGG ATC AGG GAA ACC ATG AAG AGG AAC
 C H Y L E W I R E T M K R N
 (TGA)CTATTCTAGCCTATCTCCTGTGCCCCTGACTGAGCAGAAGCCCCACAGCTGGCCAGCAGCCC

Figura 1 (continuación)

CGCCTGACATGGAACAGAACGGAGCCATCCCCAAGACCCTGTCCAAGGCCAGATGTTAGCCAAGG
ACTTGTCCACCTGAGGACAAAGCTGGCGCTCAAGGTCACCTGTTTAATGCCAAGATAACAAAGCGC
TGATCCAAGTTGCTCTGTAGGAATTTCTGTGACTTTTTCTGGGGTCAAAGAGAAACCCGAGACAC
TGTACACTGTTCTTTTCACCCACCACCCGATCCCTAGGTGAGGAGAAGCGGCTTGAAGCAGGGCT
CCATTCAATTCAACACACATGACCACCCGTGTGATCTTGAACAAGAGGCCCAATCTCACTTCGCCTTG
GTTTCCTTATCTGTAAATGAGACCATCTTATTGCTGACTTCAAAGGGCTGTTGTGAGGATTAAATG
AGATGATTCGTCTGAACTGATTAAATCGTGTCTGGCACTGA

Figura 2

zima	1	-----KKKIDMV-----VSLIAAWA---EEQ-----
KLK-L4	1	-----VWFLAL-----VLASPTLALSGGVSQES-----
KLK-L6	1	-----VSLILT-----ALQVLATAMTOS-QEDE-----
TLSP	1	-----MQRLRWLRDWKSSGRGTTAAKEPGARSSPLQAMRILQLILALATGL-VGGE-----
KLK-L3	1	-----MK-----LGLICA-----LLSLASHG---WAD-----
KLK15	1	-----VWFLILT-----ESFULASTAQ---DG-----
NES1	1	MRAPHLHLSAASGARALAKILP-----LLMAQLWAAEAL-LPQN-----
KLK-L5	1	-----MGLSTFLL-----LCVLGSQAATPK-IFN-----
neuropšina	1	-----MGRPRPRAAKTWMFY-----LLGGAWAGHSRA-QEDK-----
PSA	1	-----VWFLV-----FLTISVTWIGAAPLIL-----
hK2	1	-----VADIVL-----SIALSVSCTGAVPLIQ-----
hK1	1	-----VWFLVL-----CLAISLSGTGAAPPIQ-----
KLK-L2	1	-----MATARPPVWVICA-----LITALLGGVTEHVLANNVDVSCDHPNTPVPSGNSQDL-----
prostasa	1	-----MATAGNPWCFILG-----YUULGVAGSLVSG-----
HSCCE	1	-----MARSLLLPLOTILIL-----SUALETAGEEAQG-----
		↑
zima	20	-----NKLIVHGG-----PDKTSHEFYQAALYTSG--HLLCGGVLEHSLAVLTAAHCKRP-----
KLK-L4	24	-----EKVLNTNGTSGFLPGSYTQPPHSQPWQAALLVQG--RLLCGGVLVHSLAVLTAAHCKLE-----
KLK-L6	23	-----NKLIGGH-----TCTRSSQPWQAALLAGPRRRFLCGGALSSGQWVLTAAHCKRP-----
TLSP	52	-----TRLIK-----CFECKPHSQPWQAALFEKT--RLLCCATLASALLTAAHCKLP-----
KLK-L3	21	-----TKATG-----AEECRPNSQPWQAALFHLT--RUFCCATLISDRHLLTAAHCKRP-----
KLK15	20	-----DKYLE-----GDSQAPHSQPWQVALYERG--RFNCCASLISPHWVLSAAHCQSR-----
NES1	40	-----DTRLDEPAY-----GAPCARGSQPWQVSLFENGL--SFHCAGVLVDQSVLTAAHCGNK-----
KLK-L5	25	-----VLG-----GTECRNSQPWQVCFEGT--SLRCGGVLIDHRHVLTAAHCGSGS-----
neuropšina	33	-----VILG-----GHSQPHSQPWQAALFQGO--QLLCGGVLVGGHVLTAAHCKRP-----
PSA	23	-----SRIVG-----GWECEKHSQPWQVLVASRG--RAVCGGVLVHSLAVLTAAHCKRN-----
hK2	23	-----SRIVG-----GWECEKHSQPWQVAVYSHG--WAHCGGVLVHSLAVLTAAHCKLK-----
hK1	23	-----SRIVG-----GWECEKHSQPWQAALYHFS--TFQCGGILVHRCVLTAAHCKTSD-----
KLK-L2	51	-----GAGAGEDARSDDSSRIIN-----CSQCDMHTQPWQAALLLRP--NOLYCCAVLVHSLAVLTAAHCKKK-----
prostasa	27	-----SCQIDN-----GEGQSPHSQPWQAALVME--NEUFGSGVLVHSLAVLTAAHCFQN-----
HSCCE	28	-----DKITD-----GAPCARGSPHQAALLSG--NOLHCGGVLVHSLAVLTAAHCKNN-----
		↑
zima	67	NLQVFUGGKINR-QRESSQEQSSVRAVLIHPVDAAS-----HQQDIMLLRLARPEK--RELEFCAP-----
KLK-L4	81	GLKVYUGKHAIG-RVSGEDVREVVHSIPHEPVRRSPT-----HL--NRDHIMLLRLQSFVQ--FNGYRQLEP-----
KLK-L6	72	ILOVALGKGNR-RWEATQQLRVVROVTHEMNSRT-----HNDLMLLQLQOPAR--EGRAVRELE-----
TLSP	99	RYIVHLGQINLQ-KEEGCEQTRTATESFPHFGNNSLP-----NK--DHRNDIMLVKMASTVS--ETWAVREIT-----
KLK-L3	68	YLWVRUGGHHW-KWEGPEOLFRTVDFPHFGFNKDLN-----AN--DHRNDIMLIRPQCAR--HSPAVQEPN-----
KLK15	67	FMRVRLGPHNR-KRDGPEOLRTTSRVIPHFRY-----EA--R--SHRNDIMLLRLVQPAR--LNPQVREAV-----
NES1	91	PLWARVGDHLL-LLOG-EQLRRTTRSVVHPKIQHSGSPILPRR-----TDEHDLMLLKARPVV-PGPRVRAIQ-----
KLK-L5	67	RYWVRUGGHSLS-QLDWTEDIRHSGFSVTHPSGLGAS-----T--SHEHDLRLRLRPLVR--VTSSVQPLP-----
neuropšina	78	KYTVRLGDHSUQ-NKGGPECEIHVQSIPIHEGNNSSD-----VE--DHRNDIMLLRLQDQAS--LGSKVKEIS-----
PSA	70	KSVILGRHSLE-HPEDTCGVYQVSHSFPHPFLMDMSLLKNRFLRPGDSSHDMLRLSEPAE--LTDAVQVMD-----
hK2	70	NSQVWLGRHNL-EPEDTCGRVVSHSFPHPFLNMSLLKHQSLRPDESSHDMLRLSEPAK--RTDVVQVVG-----
hK1	70	NYQVWLGRHNL-DDENTAQFVHVSFPHFGNMSLLENKTRQADEDYSHDLMLRLTEPADTITDAVQVVE-----
KLK-L2	113	VFRVRLGHYSLSPVVYSGQGMFGGVKSIPIHFGVSHPG-----HNDLMLIKNNRIH--PTKDVREIN-----
prostasa	76	SYTIGLGLSLEADQPGSQMVEASLSVRHPEVNRPL-----LANDLMLIKIDESVS--ESDTERSES-----
HSCCE	75	EYTVHLGSDTLG--DRRACRIKASKSFRHPCSTOT-----HNDLMLIKNNRIH--LSSMVKCYR-----
		↑
zima	127	ERDSANT--SCHILGCKTA-----DG---DFEDTIQCAYTHLVREBECEHAYPGQITQNMLCAGDE-----
KLK-L4	145	LSHNNRLTPGTCRVSGWCTTT-----SE--QVNYEKLQCANIQLRSDDECRQVYPGKIDNMLCAGTK-----
KLK-L6	132	VTQACASFG--ISCRVSGWCTIS-----SE--IARYEASLOCVNINISFDEVGQKAYPRTITPGVYVAGVP-----
TLSP	163	ISSRQVTAG--ISCLISGWGCTS-----SE--QLRLPHTURCANITIIHQKGENAVPGNITDTMVCASVQ-----
KLK-L3	132	LSQTCVSPG--MQCLISGWCAVS-----SE--KALFVTLQCANISILENKLOHMAYPGHTISDSMLCAGLW-----
KLK15	127	PTTRCPHRG--EACVVSNGLYSHNEPGTAGSRSQVSLGDTURCANISIIHSDTSQDKSYPGRLTNTMVCAGAS-----
NES1	158	UPYRCAQPG--DQCVVSGWCTA-----AR--RVKYNKGLTCSSITILSPKECEVFYPGVVTNNMLCAGL-----
KLK-L5	129	IPNDQATAG--TECHVSGWCTTN-----HF--RNPFEDLQCLNLSIVSHATCHGVYPGRLTNNMVCAGG-----
neuropšina	141	ADIQTQPG--QKCVSNGWCTVT-----SE--RENFEDTLCAEVKIFPQKKCEDAYPGQITGVYVAGS-----
PSA	141	UPTQEPALG--ITCYASNGWSIE-----G--EEFLTEKKLQCVDLHVISNDVCAQVHPEQKVTKFMVCAGRW-----
hK2	141	UPTQEPALG--ITCYASNGWSIE-----G--EEFLRERSLOCVSLHLLSNDHCAQVSEKVTFFYVAGLW-----
hK1	142	UPTEEPVEG--STCLASNGWSIE-----G--ENFSFDDLCQVDLKLFPNDBCKKQHVQKVDFYVAGL-----
KLK-L2	174	VSSHQPSAG--IKCLVSGWCTTK-----SE--QVHFEKVLQCNHISVLSQKRCEDAYPGQITDINFCAG--D-----
prostasa	137	IASQOPTAG--NSCLVSGWGLLA-----NG---RMETVLCQVNVSVVSFFVGSKLIDPLYIHPSMVCAGG-----
HSCCE	133	PSRCEPFG--ITCVVSGWCTTT-----SE--DVTFESDNLQVDVKLISPODCTKVYKULLENSMLCAGIP-----

Figura 2 (continuación)

zima	187	KY	GK	DSC	Q	G	D	S	G	G	P	L	V	C	R	G	D	H	R	G	L	V	S	W	G	-	N	I	P	C	G	S	K	E	K	P	G	V	T	N	I	M	C	R	T	N	I	K	T	N	I	Q	A	K	-														
KLK-L4	208	E	G	K	D	S	C	E	G	D	S	G	G	P	L	V	C	R	N	R	T	L	Y	G	I	V	S	W	G	-	D	F	E	C	G	Q	E	D	P	G	V	T	R	S	R	G	V	L	I	W	R	E	T	I	R	K	Y	E	T	Q	Q	K	W	L	K	G	P	Q	-
KLK-L6	194	Q	G	K	D	S	C	Q	G	D	S	G	G	P	L	V	C	R	G	L	O	G	E	L	V	S	W	G	-	M	E	R	C	A	L	E	G	Y	P	G	V	T	N	I	K	R	S	N	I	E	E	T	M	R	D	K	-												
TLSP	225	E	G	K	D	S	C	Q	G	D	S	G	G	P	L	V	C	R	N	Q	S	L	O	G	T	I	S	W	G	-	Q	D	E	A	I	T	R	K	G	P	V	T	N	I	K	V	C	R	Y	L	O	I	T	E	M	K	N	-											
KLK-L3	194	E	G	R	G	S	C	Q	G	D	S	G	G	P	L	V	C	R	N	G	I	L	A	G	V	V	S	G	-	A	P	C	S	R	R	R	P	A	V	T	S	V	C	H	L	D	N	I	O	E	I	M	N	-															
KLK15	199	G	R	C	A	S	E	C	G	D	S	G	G	P	L	V	C	G	I	L	O	G	E	L	V	S	W	G	-	D	V	P	C	D	N	T	T	K	G	V	T	N	I	K	C	H	Y	I	E	N	I	R	E	T	M	R	K	-											
NES1	219	D	R	G	D	E	F	C	S	Q	G	D	S	G	P	L	V	C	D	E	T	L	O	G	E	L	S	W	G	-	V	P	C	G	S	A	O	H	S	A	V	T	O	Q	R	V	M	S	A	N	K	V	I	R	S	N	-												
KLK-L5	190	V	F	C	D	A	C	Q	G	D	S	G	G	P	L	V	C	R	G	V	L	O	G	E	L	V	S	W	G	-	S	V	G	P	C	G	D	G	I	P	G	V	T	N	I	K	V	C	H	L	D	N	I	R	M	I	R	N	-										
neuropisina	202	S	K	G	A	D	I	C	Q	G	D	S	G	G	P	L	V	C	D	E	A	L	O	G	E	L	S	W	G	-	S	D	P	C	S	R	S	D	K	P	G	V	T	N	I	K	R	C	H	Y	I	E	N	I	K	T	I	G	S	G	-								
PSA	203	T	G	G	K	D	I	C	Q	G	D	S	G	G	P	L	V	C	R	N	G	V	L	O	G	E	L	S	W	G	-	S	E	P	C	A	L	P	E	R	E	S	T	N	I	K	V	H	Y	R	K	N	I	K	T	I	V	A	N	P	-								
hK2	203	T	G	G	K	D	I	C	Q	G	D	S	G	G	P	L	V	C	R	N	G	V	L	O	G	E	L	S	W	G	-	P	E	P	C	A	L	P	E	R	E	S	T	N	I	K	V	H	Y	R	K	N	I	K	T	I	V	A	N	P	-								
hK1	204	E	G	G	K	D	I	C	Q	G	D	S	G	G	P	L	V	C	R	N	G	V	L	O	G	E	L	S	W	G	-	Y	V	P	C	E	T	N	K	S	M	A	V	R	L	S	V	Y	K	H	I	E	N	I	K	T	I	G	S	G	-								
KLK-L2	235	K	A	R	G	S	D	S	G	G	P	V	C	R	N	G	S	L	O	G	E	L	V	S	W	G	-	D	E	Y	C	A	R	E	N	R	E	P	G	V	T	N	I	K	C	H	Y	I	E	N	I	K	T</																

Figura 3

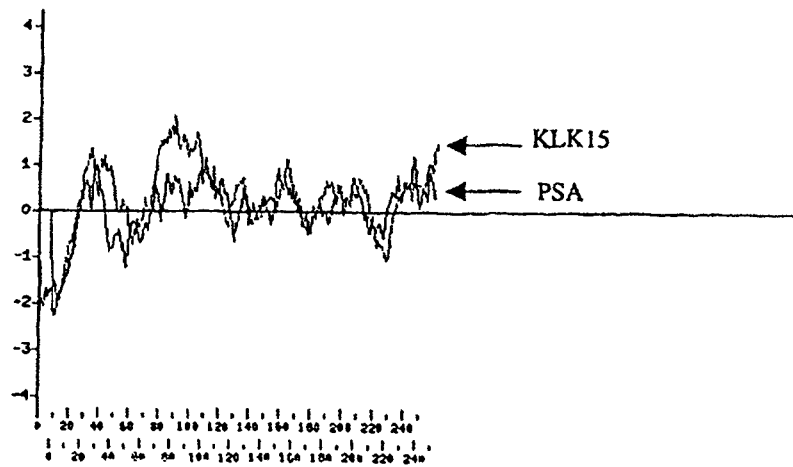


Figura 4

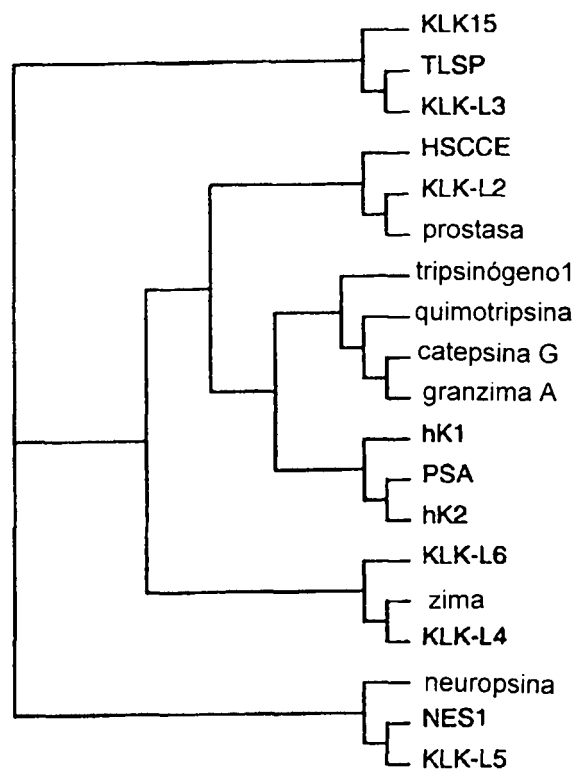


Figura 5

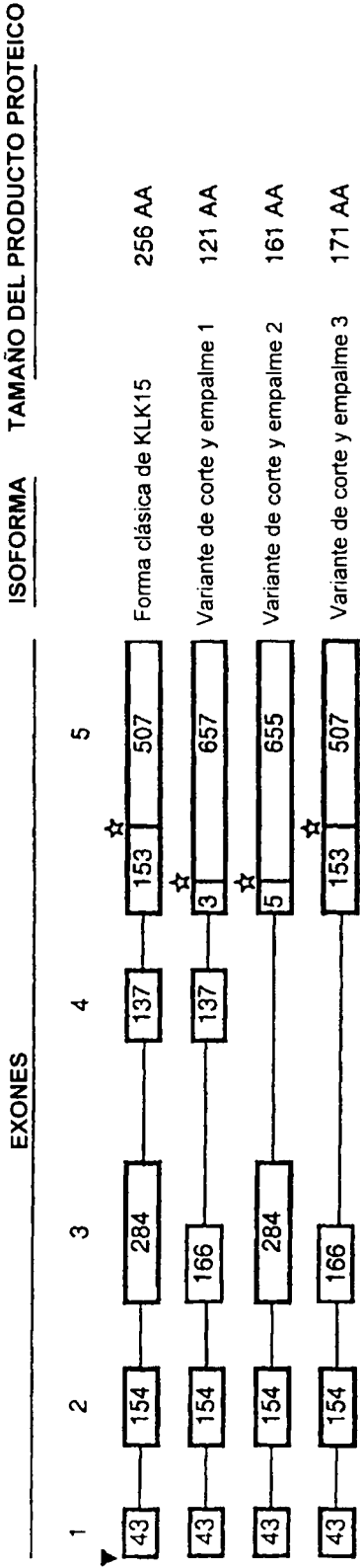


Figura 6

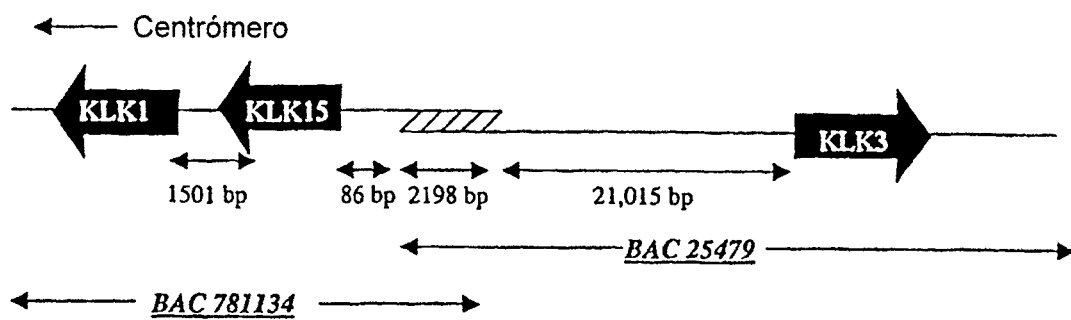


Figura 7

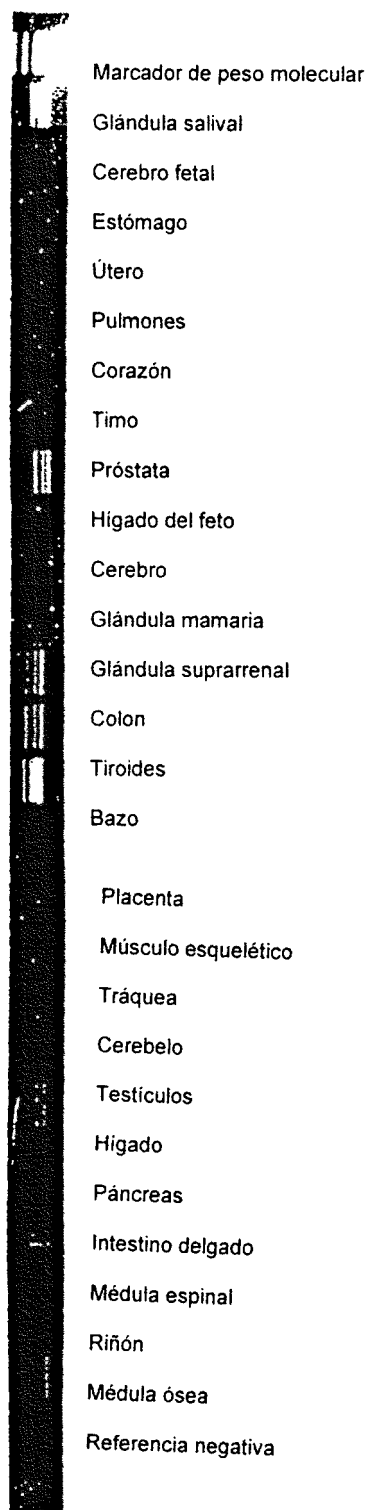


Figura 8

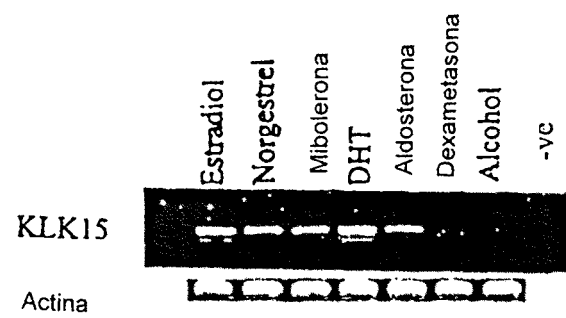
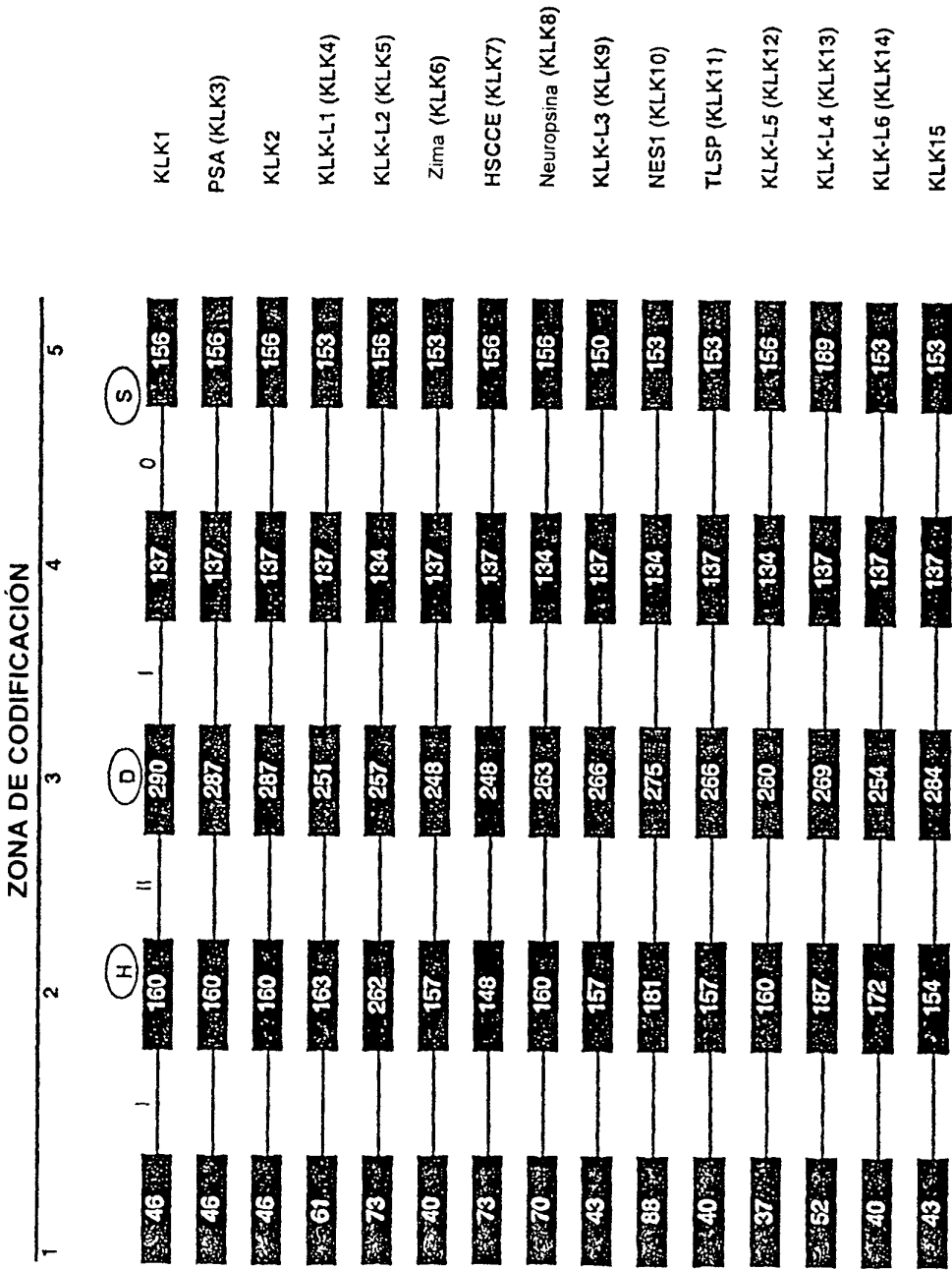


Figura 9



ES 2 266 245 T3

LISTA DE SECUENCIAS

SEC. ID nº: 1

Estructura genómica completa de KLK15

5

10

15

20

25

30

35

40

45

50

55

60

65

```

1  agaatgggtg ctgtgggatt caggggagac acctgttagg tgttggggcc tcccagaaga
61  ggtgggggca gagtgtcaga ggacaaagat gaatttggaa gatatgggga agaaggattt
121 caattcacc ctaaagcttc ctgaggcctc ccgtgggtcg ggccctgcag tactggagac
181 ccagagtga gtcagaccag ctctcgggg agctgccagt ctcgtagggg aggcagacac
241 cactgaggg caggggaggt cagagaaggc ctcaaggagg aagcggggct ggaagggaat
301 ggcgttgga atgcggtggg aggaatagcc taagcatgaa atggcaggag ggaatggc
361 agcactggc gcgtctagga caaggtcatg ggagaccag ggagaggggc tggagggaa
421 gaagccact ttgtccttga aagtgaggct ggagccaggc aactcatgcc tgtaatcca
481 gcactttggg aggctgaggc gggtagatca ctagaggtea ggagttcaag accagcctgg
541 ccaacatgg gaaactccg ctctactaaa attacaaaaa ttagctgggc gtggtggcac
601 acacctgta tcccaattgc ttgggaggct gaggcaggag aatctcttga acccagaagg
661 cagaggttac agtgagcga gatcacgcca ctccactcca acctgggcta cagagccaga
721 ctccgtctca aaaaaaaaaa aaaaaaagaa aaaaaaagaa agaaagtga tttgaagagc
781 tggactttat cctggtggtg ccaaggatcc atggagggtg gtgagcaggg gaggggcaca
841 gccagctcca gatgtagaaa gaccctttgg ggtcatggct ggagggcaag ctggtggagg
901 ggactggact ggagggggac ccaaaaggcc agataagagg gttgagatag accaggcgcg
961 gtggtcatg cctgtaatcc cagcactttg ggaggccgag gtgggtggat catgaagtca

```

ES 2 266 245 T3

1021 agagattgag gccatcctgg ctaacacggt gaaaccctgt ctctacttaa aaaaaaaaaa
 1081 tttccaaaaa attagccggg cacggtgggt ggcgcctgta gtcccagcta ctggggaggc
 1141 tgaggcggga gaatggtgtg aacctgggag gtggagcttg cagttagccg acattgtgcc
 1201 actgcactcc agcctgggtg acagagttag actccgtctc aaaaaataa aaaaagtgtg
 1261 gacagggggt ccttgctgta tgatggagag agatccaccc gctggttagca tgggtgctgga
 1321 ggctgacagg tggaggaggt ggggcagggt ctgtccgagt gcctagagga agagtaaac
 1381 ttccagagat gggggaccca gaaggaagcg cagagtgggg ttgggggaag gggataccgg
 1441 tggtcagaag aaatttatta acagtggatg ggataagtct gtgtctggag ggatcctggt
 1501 ggaggcagaa gggtcctgcc tcacctggat tctctcactc cctccccaga ctgcagccga
 1561 accctgggtcc ctctccaca atgtggcttc tctcactctc ctcttctctg ctggcatcca
 1621 cagggtagggt gggcccagga gggggcagg tctgtgggag cagggtgccc ctcccaagc
 1681 atgtctgggc ccagtgatct gccagcccct acctcaccca gagaccacta aagatccttc
 1741 ctccaccctc cacctgtgcc aatgtcccta agcccttacc gtccaggtgct ggtgtgctctg
 1801 ctctggagtc gctatgttgc ctggggcctc tcgtgcccc cgaacaggaa caggtcctg
 1861 gggttacaca aacctgagct gagtccctgg gcaaccgctt ccttgctgtg gtcctctag
 1921 ggaactgctt cacctctctg ggcttcgaat gccttctcta taagacagca cccacttgag
 1981 acaataacag tgagggtctca atagcataac agaggttaata tacatagcaa gcattagaca
 2041 agtgcctgaga ggccaacagc acagacagac tccagcttga gtcccacacc tggcactccc
 2101 cagctgcttac aggggtcttg aggggattaa atgtggttgt gtgtgaggca gaagcaaac
 2161 cctggcccag gtatgtcccc ttcagggtgtg caagccaggc acggtgctta gagcttacat
 2221 acaacgtcta tgtgtgtgtg gcaccaccga cctcatctga caaggggaag ggctgtggct
 2281 cagagggacg gccacaacat caaggtcacc ttgggtgtca ggcaaacctc agattgaact
 2341 cattctgtgt aggtgtgtgt acaccaagaa attaatgtga acctgatgcc tctctctctg
 2401 ggggtgactt tcattaacgt tctgccacaa atgacctca ctctggggg cccctgagac
 2461 cccacgcct ccagcctccc ctccggtct ctctgtgac tcacctacct gctcgcgccc
 2521 tgctgtctgc gccagctgg ggcctccacc tctctctggc ttggactggc caggtgcagc
 2581 ctctgggtccc agctgttcag ccgctaccct ccgctctcg gaggacgacc tcaccttcc
 2641 tttgttaagc ccttctgcca ccacatccgc attccctgg tctcagggg gccttggcc
 2701 cagtccctga ctgtgatggg gagagtgtgg gcatttggtc tggtgtgca aatcctgccc
 2761 ctgtgtgggt gggaggtgtc atggcttcaa ccttcagggg atgcatccac attgcccagt
 2821 ggagaggggt cctgggtcctg tgacctgaa tgtctctaat catgtcctta agcataatgc
 2881 cattctgtgt gtgtgtgtgt gtgtgtgtgt gtacatgcac gtgtgcagtg ggtatacaag
 2941 gccctgtatg ttcacatcct ctccacatgc atgagccaga tccccatatg tgaaacccaa
 3001 tcagtgaact cacagatctg gcttgggggc tgatctagag atggataaat atgtcctgcc
 3061 ctggtgctgc ctggcttcag ctgcatgtct ttgacctga atggccagcc ccgtgtctgg
 3121 gtgctgcccc agacagcaag tccacatctg agtgttggcc tctgggttg gtgtctgcag
 3181 ctctaactct acaaaatgtc ttgtgggtga atcacgggtt taaccttgac tttttttgt
 3241 ttgtttgggt ttttttgaga cggagtctcg ctctgccgcc caagctggag ttcagtgggtg
 3301 caacctcagc tcaactcaac ctccgcctcc caggttcaag caattctgtc tctgcctccc
 3361 agttagctag aattacaggc acgcaccacc acgcccagct gatttttcta ttttatttta
 3421 tllalttatt tatrrtttag tagagacggg atrtcaagat gtlggccagg crgglclcaa
 3481 actcctgacc tcaggtgalt caccacctc ggcttggcc tcccaaagtg ctgggaltac
 3541 aggcgtgagc caccacacct gggcaacctt gactatttat tataggtaat lctgtgcaga
 3601 tgtctgactt atgttggcca tctccaggat ggacctgaac tttcnacagt atgtcctgtl
 3661 gactaaatcc aggtgtcatt tgcaaaaaac aactaatatt attaagtgc taccagggtc
 3721 aggtatcact caccatacat acacacatgc acacacacac atacacattc ctacctcatc
 3781 cttacaacaa tcttcatttt acagatgagg aaacagaggc acagacaggt cgaataactt
 3841 actcaaagtt tcacagctag tacattcgaa ccaggctta aggaccatc tttgtccaga
 3901 ccctgtatgc aagtgtctgt gacactggat gccaaagctc acactagaga tgttgaattt
 3961 aggtctgaac aatatccaat tctgtgtgtg tgtttgtgtg tgcatgtgtg tgtgtgtatg
 4021 tattcatgtc ttaaccatcc atattcatat acacatatga acatctgtgc tgtgattctt
 4081 tttttttttt tttttttttt tttgagatgg agtttctact ttgtcaccga ggctggagtg
 4141 caatggagca acctccgctc actgcaatct ccgctcccg ggttcaagcg attttcctgc
 4201 ctacgctcc agagttagct ggattacagg caccgcccac catgcccagc taattttttg
 4261 tttttttgtt agagacagtg tttccccata ttggccaggc ttgtctcgaa ctccttgact
 4321 cagggtgatcc acccgctcgc gctcccaaaa gtgtgggat tacaggcatg agccaccgtg
 4381 cccagcctgt gctgtgattc ttgaagctgc aacctatgtg catgcaagtg aatttcagct
 4441 tccagctcctg tccatagctg tacetaagtg ttggaagctg atgtgcatgt atgcatgtcc
 4501 atgacctgt atagccacat ctgggactca tactgcacac tgaatttggc tgacatgtcc
 4561 agactctggg gccaaaggct ggacacacat actgagtggc cacatgcgtt tgacgtctgt
 4621 gacaatttgg tgaccgtgaa tgactggttt caagtgaaca cctgtctgaa cctgtatcca
 4681 gtgcccctgt ctccaccccc aaccacagag gacttcttgc cctctggtct gttcccttc
 4741 ctctctctcc cagagtctta tagcaaatgg ggtgggggct agagtctctg agaaaacagg

ES 2 266 245 T3

4801 cagcgggtgt aaataaacaa cagggcagggc ggagcatggt ggctcacacc tgtaatceca
 4861 gcactttggg agggctgaggc gggcagagca tttgaagtca gaagtttgag actacctggc
 4921 taacatggtg agacctcgct tctactaaaa atacaaaaat tagccagggt tgggtggcggg
 4981 cacctcagct actcgggagg ctgaggcagg aggatcactt gaaccaggga ggcggaagtt
 5041 gcagttagct gagatcatgc cactgcactc cagcctgggc aaaagagtga gactccgtct
 5101 caaaaacaac aacaacaaca aaacaaaaaa cagggcaggg tgtcttgaga agttagggga
 5161 aaggcatagg catatagtag tttagggcagg gtgcaaggaa ggtgtaggag gcaatgtaaa
 5221 cgtccctgtc ctcaggcatc ctctaccctt tctcttagca gcccaggatg gtgacaagtt
 5281 gctggaaggt gacgagtgtg caccctcact ccagccatgg caagtggctc tctacgagcg
 5341 tggacgcttt aactgtggcg ctccctcatc ctccccacac tgggtgctgt ctgcgcccca
 5401 ctgcccagaag cggtatgaag gcaggggctc agggctcctga gggagcctgg ttcgggggga
 5461 agagctccta gatttggggg aagacggagg cagacgccag aactcctggg tctgaaagga
 5521 cgaggaggcc ggatgtcaag ccctggggtt aggaaggagt gtgtgtttca aagccttcga
 5581 tctctgaagg aggaaggaga agactagtct cagcttttga gcctcagttc tagggatgtg
 5641 agaactcctg attcggggac agaccaggag ggggctggga gtagttggag gggatcgagt
 5701 tctaggagtg tgcctgactt cagactcggt ggtccttgag gagcaggggc tgggaaccatt
 5761 ggcttcaggg tcttgggaaa aggtaattgg atgtcgagat ttctaaaggg tcgggagacc
 5821 tcgggttgcc cactctttga tctttctgtc ctctacttgc gggtaaccac tggcccgcac
 5881 tccactggcg ggaaaaccac tcgcccgcac agcttcatga gagtgcgcac gggagagcac
 5941 aacctgcgca agcgcgatgg ccagagcaac ctacggacca cgtctcgggt cattccacac
 6001 ccgcgctacg aagcgcgcag ccaccgcaac gacatcatgt tgctgcgcct agtccagccc
 6061 gcacgcctga acccccagggt gcgcccgcg gtgtaccaca cgcgttgccc ccaccggggg
 6121 gaggcctgtg tgggtgtctg ctggggcctg gtgtcccaac acgagcctgg gaccgctggg
 6181 agcccccgtg cacaagtgct gtgaaaggat ggagctggat gcgaggcctg aaggaatcct
 6241 atgtctcagg gctcttgggc ggaggggaca agggccggaa tttatggatc tgctccaagt
 6301 ccactgtctt ccccgatgag tctcccagat acgttgcatl gtgccaacat cagcatatct
 6361 tcggacacat cttgtgacaa gagctaccca gggcgcttga caaacaccat ggtgtgtgca
 6421 ggcgcggagg gcagagcgcg agaatcctgt gaggtcagag cctagagggg ccatacaggc
 6481 gaagaagagg gatggggaca ggtgtgggag tccggatggg gttggatttt ctttctttg
 6541 ggccagagaa gatgttaggg tttaggttgg agatggagta ggaagagaa ttagaatagg
 6601 ggtgaggttg gagttggggt tataggtggg gattgcgttg tttgaggttg ataactgtga
 6661 tgggttggct gagatggcat gggttggggt tgagaattgg aatggtttgg tttgatctct
 6721 ggtgggaaat acgtcagggt tgaattggga tgaggtagat tttgtttgga atgcagaaga
 6781 catgaagatt gagattggat tttgagatgg gcatgggttt gatttgattt tgaatgggtga
 6841 ggaatgtggc tgagttggat ttaacttagt acagttgcac tggagttgca tgggggtgag
 6901 attggatata ggttgggtga gttgtattga gctgtgttga attggggttg ggggtggggt
 6961 tgggttggct ctgtttggga taaactgggc tgatttgagt tgagttgggt tggggttccc
 7021 tgggatgggg atggattggg tttggggtga gattgcaaat ggtgattagg atgaggatga
 7081 atccaggagg tttactcaa cctgagacc cctcttttcc ccacagggtg actctggggg
 7141 acccctggtc tgtgggggca tcctgcaggg catttgttcc tgggggtgacg tcccctgtga
 7201 aacaccacc aagcctgggt tctataccaa agtctgccac tacttgagtg ggatcaggga
 7261 aaccatgaag aggaactgac tattctagcc tatctcctgt gccctgactg gagcagaagc
 7321 cccacagctg gccacgagc ccgcctgac atggaacaga acggagccat cccccaagac
 7381 cctgtccaag gccacagatg tagccaagga cttgtcccaac ctgagyaaca agclggcgct
 7441 aggggttcacc lgtttaatgc caagataaca aagcgttgat ccaagtttgc clgtaggaa
 7501 tttctgtgact tttttctggg gtcaaaagaga aaccccgaga cactgtacac tgttcccttt
 7561 caccaccacc cccgatccct aggtgaggag aagcggcltg aagcagggtc ccattcaltc
 7621 aacacacatg accaccctg lgtatcttgaa caagaggccc aatctcactt cgccllggtt
 7681 tcttatctctg taaaatgaga ccatcttatt gctgacttca aagggtctgt gtgaggatta
 7741 aatgagatga ttcgtctgaa ctgattaaaa tctgtgtctg cactgagtaa ataccctcta
 7801 tctctggatc ccagttaaag gacctaacag acactagatt accaagaatg gctttttctt
 7861 taagggtttg tctcgggctg ggcattggtg ctacacctg taatcccagc actttgggag
 7921 gccaaaggcg gcggctcact tgaggtcagg agtgcaagac cagcctggcc aacatggaga
 7981 aaccccatct ctactaaaaa tactaaaaaa atttagccgg gcgtggtggc acacgactgt
 8041 aatccttagct acttgggagg gtgatgtggg aggatcgctt gaacttagga ggcaggagt
 8101 gcagttagcc gagatcgcc cactgcactc cagcctgggt acagagcaag actccatctc
 8161 agaaaaaa aaaaaaa aagatttag tttcgggctt cctggtagcc atggcaaaaa
 8221 ggcaaatact gtcctttcct tagccagggt cctgatatac agcagagggt ggaactctga
 8281 gctgctttga ttttaccaaa aagccaagac aacctgttgg aagcctatgg gtttaccatt
 8341 gaggtgagc gaatctagtt cctaattatc ttcagagacc acaaaatgtg atgttcaagg
 8401 tcgctgaatg ttgaagtaca tgaacctggc tctgtgagacc taaatattgt actggtggtg
 8461 ggggggaagg gtcattggaa tctgtgttga gcttgatctt gacctgcgag ggaaggttgt
 8521 ccagatctct ggacttttga ggaccgagct tgagcaccat aatgggagca gaagtgcgag
 8581 gtcctttgaga cccgcttctg tggggcggcg ccggtatttg atgctaaaaa ttacctggga
 8641 accctgaata catctgggtt gggcgacaaa tgtgtggctc cccacacatc ttttaggaaca
 8701 catttgggca acccggtggg agtgaacggc ctggc

ES 2 266 245 T3

SEC. ID nº: 2

ARNm clásico (1581..1623, 5259..5412, 5913..6196, 6317, 6453)

5 ATGTGGCTTCTCCTCACTCTCTCCTTCTGCTGGCATCCACAGCAGCCAGGATGGTGACAAGTTGCTGGAAGGTG
ACGAGTGTGCACCCCACTCCCAGCCATGGCAAGTGGCTCTCTACGAGCGTGGACGCTTTAACTGTGGCGCTTCCCT
CATCTCCCCACACTGGGTGCTGTCTGCGGCCCACTGCCAAAGCCGCTTCATGAGAGTGCCTGGGAGAGCACAAC
CTGCGCAAGCGCGATGGCCAGAGCAACTACGGACCACGTCTCGGGTCATTCCACACCCGCGCTACGAAGCGCGCA
10 GCCACCGCAACGACATCATGTTGCTGCGCCTAGTCCAGCCCGCACGCTGAACCCCAAGGTGCGCCCCGCGGTGCT
ACCCACGCGTTGCCCCACCCGGGGAGGCCTGTGTGGTGTCTGGCTGGGGCCCTGGTGTCCCAACAGAGCCTGGG
ACCGCTGGGAGCCCCGGTCACAAGTGAGTCTCCAGATACGTTGCATTGTGCCAACATCAGCATTATCTCGGACA
CATCTTGTGACAAGAGCTACCCAGGGCGCTGACAAACACCATGGTGTGTGCAGGCGCGAGGGCAGAGCGCAGA
ATCCTGTGAGGGTGACTCTGGGGGACCCCTGGTCTGTGGGGGCATCTGCAGGGCATTGTCTCTGGGTGACGTC
15 CCTTGTGACAACACCACCAAGCCTGGTGTCTATACCAAGTCTGCCACTACTTGGAGTGGATCAGGGAAACCATGA
AGAGGAAC TGACTATTCTAGCCTATCTCCTGTGCCCCCTGACTGAGCAGAAGCCCCACAGCTGGCCAGCAGCCCCG
CCTGACATGGAACAGAACGGAGCCATCCCCCAAGACCCCTGTCCAAGGCCAGATGTTAGCCAAGGACTTGTCCAC
CTGAGGACAAAGCTGGCGCTCAAGGTCACTGTTTAAATGCCAAGATAACAAAGCGCTGATCCAAGTTGCTCTGTAG
GAATTTCTGTGACTTTTTTCTGGGGTCAAAGAGAAACCCGAGACACTGTACACTGTTCTTTTCAACCCACCAACC
20 CGATCCCTAGGTGAGGAGAAGCGGCTTGAAGCAGGGCTCCATTCAATCAACACACATGACCACCCGTGTGATCTTG
AACAAGAGGCCCAATCTCACTTCGCTTGGTTCTTATCTGTAATAATGAGACCATCTTATTGCTGACTTCAAAGG
GCTGTGTGAGGATTAAATGAGATGATTCTGCTGAACGTGATTAATAATCGTGTCTGGCACTGA

SEC. ID nº: 3

Estructura de la VARIANTE 1 DE CORTE Y EMPALME DE ARNm de KLK 15 (1581..1623, 5259..5412, 5913..6078, 6317..6453)

25 ATGTGGCTTCTCCTCACTCTCTCCTTCTGCTGGCATCCACAGCAGCCAGGATGGTGACAAGTTGCTGGAAGGTG
ACGAGTGTGCACCCCACTCCCAGCCATGGCAAGTGGCTCTCTACGAGCGTGGACGCTTTAACTGTGGCGCTTCCCT
CATCTCCCCACACTGGGTGCTGTCTGCGGCCCACTGCCAAAGCCGCTTCATGAGAGTGCCTGGGAGAGCACAAC
30 CTGCGCAAGCGCGATGGCCAGAGCAACTACGGACCACGTCTCGGGTCATTCCACACCCGCGCTACGAAGCGCGCA
GCCACCGCAACGACATCATGTTGCTGCGCCTAGTCCAGCCCGCACGCTGAACCCCAAGTGTCTCCAGATACG
TTGCATTGTGCCAACATCAGCATTATCTCGGACACATCTTGTGACAAGAGCTACCCAGGGCGCTGACAAACACCA
TGGTGTGTGCAGGCGCGGAGGGCAGAGGCGCAGAATCTGTGAGGTGACTCTGGGGGACCCCTGGTCTGTGGGGG
CATCTGACGGGCATTGTGTCTGGGGTGACGTCCCTTGTGACAACACCACCAAGCCTGGTGTCTATACCAAAGTC
35 TGCCACTACTTGGAGTGGATCAGGGAAACCATGAAGAGGAAC TGACTATTCTAGCCTATCTCCTGTGCCCTGACT
GAGCAGAAGCCCCACAGCTGGCCAGCAGCCCCGCTGACATGGAACAGAACGGAGCCATCCCCCAAGACCCCTGT
CAAGGCCAGATGTTAGCCAAGGACTTGTCCCACTGAGGACAAAGCTGGCGCTCAAGGTCACTGTTTAAATGCCA
AGATAACAAAGCGCTGATCCAAGTTGCTCTGTAGGAATTTCTGTGACTTTTTTCTGGGGTCAAAGAGAAACCCGA
GACACTGTACACTGTTCTTTTCAACCCACCAACCCGATCCCTAGGTGAGGAGAAGCGGCTTGAAGCAGGGCTCCAT
40 TCAATTCAACACACATGACCACCCGTGTGATCTTGAACAAGAGGCCCAATCTCACTTCGCTTGGTTTCTTATCTG
TAAATGAGACCATCTTATTTGCTGACTTCAAAGGGCTGTTCTGAGGATTAAATGAGATGA

SEC. ID nº: 4

Estructura 2 de CORTE Y EMPALME del ARNm de KLK 15 (1581..1623, 5259..5412, 5913..6196, 7127..7786)

45 ATGTGGCTTCTCCTCACTCTCTCCTTCTGCTGGCATCCACAGCAGCCAGGATGGTGACAAGTTGCTGGAAGGTG
ACGAGTGTGCACCCCACTCCCAGCCATGGCAAGTGGCTCTCTACGAGCGTGGACGCTTTAACTGTGGCGCTTCCCT
CATCTCCCCACACTGGGTGCTGTCTGCGGCCCACTGCCAAAGCCGCTTCATGAGAGTGCCTGGGAGAGCACAAC
50 CTGCGCAAGCGCGATGGCCAGAGCAACTACGGACCACGTCTCGGGTCATTCCACACCCGCGCTACGAAGCGCGCA
GCCACCGCAACGACATCATGTTGCTGCGCCTAGTCCAGCCCGCACGCTGAACCCCAAGTGTCTCCAGATACG
TGGTGTGTGCAGGCGCGGAGGGCAGAGGCGCAGAATCTGTGAGGTGACTCTGGGGGACCCCTGGTCTGTGGGGG
ACCCACGCGTTGCCCCACCCGGGGAGGCCTGTGTGGTGTCTGGCTGGGGCCCTGGTGTCCCAACAGAGCCTGGG
ACCGCTGGGAGCCCCCGGTACAAGGGTGACTCTGGGGGACCCCTGGTCTGTGGGGGCATCTGCAGGGCATTGTG
55 TCCTGGGGTGACGTCCCTTGTGACAACACCACCAAGCCTGGTGTCTATACCAAAGTCTGCCACTACTTGGAGTGA
TCAGGGAAACCATGAAGAGGAAC TGACTATTCTAGCCTATCTCCTGTGCCCTGACTGAGCAGAAGCCCCACAGC
TGGCCAGCAGCCCCGCTGACATGGAACAGAACGGAGCCATCCCCCAAGACCCGTGCCAAGGCCAGATGTTAGCC
AAGGACTGTGCCACCTGAGGACAAAGCTGGCGCTCAAGGTCACTGTTTAAATGCCAAGATAACAAAGCGCTGATC
CAAGTTGCTCTGTAGGAATTTCTGTGACTTTTTTCTGGGGTCAAAGAGAAACCCGAGACACTGTACACTGTTCT
60 TTTCAACCCACCAACCCGATCCCTAGGTGAGGAGAAGCGGCTTGAAGCAGGGCTCCATTCAATCAACACACATGACC
ACCCGTGTGATCTTGAACAAGAGGCCCAATCTCACTTCGCTTGGTTTCTTATCTGTAATAATGAGACCATCTTAT
TGCTGACTTCAAAGGGCTGTGTGAGGATTAAATGAGATGA

ES 2 266 245 T3

SEC. ID n°: 5

Estructura 3 de CORTE Y EMPALME del ARNm de KLK 15 (1581..1623, 5259..5412, 5913..6078, 7127..7786)

5

ATGTGGCTTCTCCTCACTCTCTCCTTCCTGCTGGCATCCACAGCAGCCCAGGATGGTGACAAGTTGCTGGAAGGTG
ACGAGTGTGCACCCCACTCCCAGCCATGGCAAGTGGCTCTCTACGAGCGTGGACGCTTTAACTGTGGCGCTTCCCT
CATCTCCCCACACTGGGTGCTGTCTGCGGCCCACTGCCAAAGCCGCTTCATGAGAGTGCCTGGGAGAGCACAAAC
CTGCGCAAGCGCGATGGCCAGAGCAACTACGGACCACGTCTCGGGTCATTCACACCCGCGCTACGAAGCGCGCA
GCCACCGCAACGACATCATGTTGCTGCGCCTAGTCCAGCCCGCACGCCCTGAACCCCCAGGGTGACTCTGGGGGACC
CCTGGTCTGTGGGGGCATCCTGCAGGGCATGTGTCTTGGGGTGACGTCCCTTGTGACAACACCACCAAGCCTGGT
GTCTATACCAAAGTCTGCCACTACTTGGAGTGGATCAGGGAACCATGAAGAGGAACTGACTATTCTAGCCTATCT
CCTGTGCCCTGACTGAGCAGAAGCCCCACAGCTGGCCAGCAGCCCCGCTGACATGGAACAGAACGGAGCCATC
CCCCAAGACCCTGTCCAAGGCCAGATGTTAGCCAAGGACTTGTCACCTGAGGACAAAGCTGGCGCTCAAGGTC
ACCTGTTTAAATGCCAAGATAACAAAGCGCTGATCCAAGTTGCTCTGTAGGAATTTCTGTGACTTTTTTCTGGGGTC
AAAGAGAAACCCCGAGACACTGTACACTGTTCCCTTTTACCCACCACCCCGATCCCTAGGTGAGGAGAAGCGGCTT
GAAGCAGGGCTCCATTTCATTCAACACACATGACCACCCGTGTGATCTTGAACAAGAGGCCCAATCTCACTTCGCCT
TGGTTTCCTTATCTGTAATAATGAGACCATCTTATTGCTGACTTCAAAGGGCTGTTGTGAGGATTAAATGAGATGA

10

15

20

SEC. ID n°: 6

Proteína de KLK15

25

MWLLLTLSFLLASTAAQDGDKLLLEGDECAHPSQPWQVALYERGRFNCGLISPHWVLSAAHCQSRFMRVRLGEHN
LRKRDGPQLRTTSRVIPHPRYEARSHRNDIMLLRLVQPARLNPQVRPAVLPTRCPHPEGEACVVSGWGLVSHNEPG
TAGSPRSQVSLPDTLHCANISIIISDTSCKSYPGRLTNTMVCAGAEGRGAESCEGDSGGPLVCGGILQGIIVSWGDV
PCDNTTKPGVYTKVCHYLEWIRETMKRN

30

SEC. ID n°: 7

Variante 1 de corte y empalme de KLK15

35

MWLLLTLSFLLASTAAQDGDKLLLEGDECAHPSQPWQVALYERGRFNCGLISPHWVLSAAHCQSRFMRVRLGEHN
LRKRDGPQLRTTSRVIPHPRYEARSHRNDIMLLRLVQPARLNPQ

40

SEC. ID n°: 8

Variante 2 de corte y empalme de KLK15

45

MWLLLTLSFLLASTAAQDGDKLLLEGDECAHPSQPWQVALYERGRFNCGLISPHWVLSAAHCQSRFMRVRLGEHN
LRKRDGPQLRTTSRVIPHPRYEARSHRNDIMLLRLVQPARLNPQVRPAVLPTRCPIIPGEACVVSGWGLVSHNEPG
TAGSPRSQ

SEC. ID n°: 9

Variante 3 de corte y empalme de KLK15

50

MWLLLTLSFLLASTAAQDGDKLLLEGDECAHPSQPWQVALYERGRFNCGLISPHWVLSAAHCQSRFMRVRLGEHN
LRKRDGPQLRTTSRVIPHPRYEARSHRNDIMLLRLVQPARLNPQGDSSGGPLVCGGILQGIIVSWGDVPCDNTTKPG
VYTKVCHYLEWIRETMKRN

55

SEC. ID n°: 10

HNEPGTAG

60

65

ES 2 266 245 T3

SEC. ID nº: 11

5' no traducida

5

1-1580

10

15

20

25

30

35

AGAATGGGTGCTGTGGGATTTCAGGGGAGACACCTGTTAGGTGTTGGGGCC
TCCCAGAAGAGGTGGGGGCAGAGTGTTCAGAGGACAAAGATGAATTTGGAA
GATATGGGGAAGAAGGATTTCAATTCACCCCTCAAAGCTTCCTGAGGCCTC
CCGTGGGTTCGGGCCCTGCAGTACTGGAGACCCAGAGTGGAGTCAGACCAG
CTCTCGGGGAGCTGCCAGTCTCGTAGGGGAGGCAGACACCACCTGAGGGT
CAGGGGAGGTCAGAGAAGGCCTCAAGGAGGAAGCGGGCTGGAAGGGAAT
GGCGTTGGATATGCGGTGGGAGGAATAGCCTAAGCATGAAATGGCAGGAG
GGAAAAATGGCAGCACTGGCTGCGTCTAGGACAAGGTCATGGGAGACCCAG
GGAGAGGGGCTGGAAGGGAAGAAGCCACTTTTGTCTTTGAAAGTGAGGCT
GGAGCCAGGCAACTCATGCTGTAAATCCCAGCACTTTGGGAGGCTGAGGC
GGGTGGATCACTAGAGGTCAGGAGTTCAAGACCAGCCTGGCCAAACATGGT
GAAACTCCGTCTCTACTAAAAATTACAAAAATTAGCTGGGCGTGGTGCCAC
ACACCTGTAAATCCCAATTGCTTGGGAGGCTGAGGCAGGAGAATCTCTTGA
ACCCAGAAGGCAGAGGTTACAGTGAGCGGAGATCACGCCACTCCACTCCA
ACCTGGGCTACAGAGCCAGACTCCGTCTCAAAAAAAAAAAAAAAAAAGAA
AAAAAAGAAAGAAAGTGAATTTGAAGAGCTGGACTTTATCCTGGTGGTG
CCAAGGATCCATGGAGGGTGGTGAGCAGGGGAGGGGCACAGCCAGCTCCA
GATGTAGAAAGACCCCTTTGGGGTCATGGCTGGAGGGCAAGCTGGTGGAGG
GGACTGGACTGGAGGGGGACCCAAAAGGCCAGATAAGAGGGTTGAGATAG
ACCAGGCGCGGTGGCTCATGCTGTAAATCCCAGCACTTTGGGAGGCGGAG
GTGGGTGGATCATCAAGTCAAGAGATTGAGGCCATCCTGGCTAACACGGT
GAAACCCGTCTCTACTTAAAAAAAAAAAAATTTCCAAAAATTAGCCGGG
CACGGTGGTGGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCGGGA
GAATGGTGTGAACCTGGGAGGTGGAGCTTGCAAGTGAGCCGACATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGTGAGACTCCGTCTCAAAAAAATAA
AAAAAGTTGGGACAGGGGGTCCCTGCGTGATGATGGAGAGAGATCCACCC
GCTGGTAGCATGGTGTGAGGCTGACAGGTGGAGGAGGTGGGGCAGGGT
CTGTCCGAGTGCCTAGAGGAAGAGTAAACCTTCCAGAGATGGGGGACCCA
GAAGGAAGCGCAGAGTGGGGTTGGGGGAAGGGGATACCGGTGCTCAGAAG
AAATTTATTAACAGTGGATGGGATAAGTCTGTGTCTGGAGGGATCCTGGT
GGAGGCAGAAGGGTCTGCTCACCTGGATTCTCTCACTCCCTCCCCAGA
CTGCAGCCGAACCCCTGGTCCCTCCTCCACA

SEC. ID nº: 12

Exón 1

1581-1623

45

ATGTGGCTTCTCCTCACTCTCTCCTTCCTGCTGGCATCCACAG

50

55

60

65

Intrón 1

5

10

15

20

25

30

35

40

45

50

Exón 2

55

Intrón 2

60

65

7

ES 2 266 245 T3

SEC. ID nº: 16

Exón 3 (variante 2 clásica y de corte y empalme)

5913-6196

cttcatga gagtgcgcct gggagagcac aacctgcgca agcgcgatgg cccagagcaa ctacggacca
cgtctcgggt cattccacac ccgcgctacg aagcgcgcag ccaccgcaac gacatcatgt tgctgcgcct
agtcacagccc gcacgcctga acccccaggt gcgccccgcg gtgctaacca cgcgttgccc ccaccgggg
gaggcctgtg tgggtgcttg ctggggcctg gtgtcccaca acgagcctgg gaccgctggg
agccccggg cacaag

SEC. ID nº: 17

Intrón 3 - Variante 2 clásica y de corte y empalme

6197-6316

tgc gtgaaaggat ggagctggat gcgaggcctc aaggaatecta tgctccaggg ctcttgggcg
gaggggaca agggccggaa ttatggatc tgctccaagtccactgtctt ccccag

SEC. ID nº: 18

Exón 3 - (variante de corte y empalme 1 y 3)

5913-6078

CTTCATGAGAGTGCCTCGGGAGAGCACAACCTGCGCAAGCGCGATGGCCAGAGCAACTACGGACCACGTCTCGG
GTCATTCCACACCCGCGCTACGAAGCGCGCAGCCACCGCAACGACATCATGTTGCTGCGCCTAGTCCAGCCCGCAC
GCCTGAACCCCAAGGTGCGCCCCGCGGTGCTACCCACGCGTTGCCCCACCCGGGGAGGCCTGTGTGGTGTCTGG
CTGGGGCCTGGTGTCCCAACAGAGCCTGGGACCGCTGGGAGCCCCCGGTCAACAAG

SEC. ID nº: 19

Intrón 3 - (variante de corte y empalme 1 y 3)

6079-6316

GTGCGTGAAGGATGGAGCTGGATGCGAGGCCCTAAGGAATCCTATGCTCCAGGGCTCTTGGGCGGAGGGGACAAG
GGCCGGAATTTATGGATCTGCTCCAAGTCCACTGTCTTCCCCAG

SEC. ID nº: 20

Exón 4

6317-6453 (variante 1 clásica y de corte y empalme)

TGAGTCTCCAGATACGTTGCATTGTGCCAACATCAGCATTTATCTCGGACACATCTTGTGACAAGAGCTACCCAGG
GCGCCTGACAAACACCATGGTGTGTGCAGGCGCGGAGGCGAGAGGCGCAGAATCCTGTGAG

SEC. ID nº: 21

Intrón 4 (variante 1 clásica y de corte y empalme)

6454-7126

GTCAGAGCCTAGAGGGCCATCAGGCGGAAGAAGAGGGATGGGGACAGGTGTGGGAGTCCGGATGGGGTTGGATTT
TCTTTGCTTTGGGCCAGAGAAGATGCTAGGGTTAGGCTTGGAGATGGAGTAGGAAGAGAAGTTAGAATAGGGGTGA
GGTTGGAGTTGGGGTTATAGGTGGGGATGCGTTGTTTGAAGTGGATAACTGTGATAGTTAGTTTGAGATGGCATTG
GGTTGGGGTTGAGAATGGGAATGGTTTGGTTTGTATCTGGGTGGGAAATACGTCAGGGTTGAATTGGGATGAGGTA
GATTTTGTTTGGAAATGCAGAAGACATGAAGATTGAGATTGGATTTTGAGATGGGCATGGGTTTGATTGATTTTGA
ATGGTGAGGATGTGGGCTGAGTTGGATTTAACTTAGTACAGTTGCACTGGAGTTGCATGGGGGTGAGATTGCATAT
AGGTTGGGTGAGTTGTATTGAGCTGTGTGAATTGGGGTTGGGGTTGGGGTTGGGTGGCTCTGTTTGGGATAAAC
TGGGCTGTATTGAGTTGAGTTGGGTGGGGTTCCCTGGGATGGGGATGGATTGGGTGGGGTGAGATTGCAAAATG
GTGATTAGGATGAGGATGAATCCAGGAGTTTCACTCAACCTGAGACCCCTCTTTTCCCCACAG

ES 2 266 245 T3

SEC. ID nº: 22

Intrón 4 (variantes 2 y 3 de corte y empalme)

6079-7126

T GCGCCCCGCG GTGCTACCCA CGCGTTGCC CCACCCGGG GAGGCCTGTG TGGTGTCTGG
CTGGGGCCTG GTGTCACACA ACGAGCCTGG GACCGCTGGG AGCCCCGGT CACAAGGTGC GTGAAAGGAT
GGAGCTGGAT GCGAGGCCTC AAGGAATCCTATGCTCCAGG GCTCTTGGGC GGAGGGGACA AGGGCCGGAA
TTTATGGATC TGCTCCAAGT CCACTGTCTT CCCCAGTGAG TCTCCAGAT ACGTTGCATT GTGCCAACAT
CAGCATTTATC TCGGACACAT CTTGTGACAA GAGCTACCCA GGGCGCCTGA CAAACACCAT GGTGTGTGCA
GGCGCGGAGG GCAGAGGCGC AGAATCCTGT GAGGTCAGAG CCTAGAGGGG CCATCAGGCG GAAGAAGAGG
ATGGGGACA GGTGTGGGAG TCCGGATGGG GTTGGATTTT CTTTGCTTTG GGCCAGAGAA GATGCTAGGG
TTAGGCTTGG AGATGGAGTA GGAAGAGAAG TTAGAATAGG GGTGAGGTTG GAGTTGGGGT TATAGGTGGG
GATTGCGTGG TTTGAGGTGG ATAACGTGA TAGTTAGTTT GAGATGGCAT GGGTTGGGGT TGAGAATGGG
AATGGTTTGG TTTGATTCTG GGTGGGAAAT ACGTCAGGGT TGAATTGGGA TGAGGTAGAT TTTGTTTGA
ATGCAGAGA CATGAAGATT GAGATTGGAT TTTGAGATGG GCATGGGTTT GATTGTGATT TGAATGGTGA
GGATGTGGGC TGAGTTGGAT TTAACTTAGT ACAGTTGCAC TGGAGTTGCA TGGGGGTGAG ATTGGATATA
GGTTGGGTGA GTTGTATTGA GCTGTGTTGA ATTGGGGTTG GGGTTGGGGT TGGGTTGGCT CTGTTTGGGA
TAAACTGGGC TGTATTGAGT TGAGTTGGGT TGGGGTTCCC TGGGATGGGG ATGGATTGGG TTTGGGGTGA
GATTGCAAAT GGTGATTAGG ATGAGGATGA ATCCAGGAGG TTCACTCAA CCTGAGACCC CCTCTTTTCC
CCACAG

SEC. ID nº: 23

Exón 5

7127-7786

GGTGACTCTGGGGACCCCTGGTCTGTGGGGCATCCTGCAGGGCATTGTGTCTGGGGTGACGTCCCTTGTGACA
ACACCACCAAGCCTGGTGTCTATACCAAAGTCTGCCACTACTTGGAGTGGATCAGGAAACCATGAAGAGGAAC TG
ACTATTCTAGCCTATCTCCTGTGCCCCGACTGAGCAGAAAGCCCCACAGCTGGCCAGCAGCCCCGCTGACATGG
AACAGAACGGAGCCATCCCCAAGACCCTGTCCAAGGCCAGATGTTAGCCAAGGACTTGTCCCACTGAGGACAA
AGCTGGCGCTCAAGGTCACCTGTTTAAATGCCAAGATAACAAAGCGCTGATCCAAGTTGCTCTGTAGGAATTTCTGT
GACTTTTCTTCTGGGGTCAAAGAGAAACCCGAGACACTGTACACTGTTCCTTTTACCCACCACCCCGATCCCTAG
GTGAGGAGAAGCGGCTTGAAGCAGGGCTCCATTCAATCAACACACATGACCACCCGTGTGATCTTGAACAAGAGGC
CCAATCTCACTTCGCTTGGTTTCTTATCTGTAAATGAGACCATCTTATTGCTGACTTCAAAGGGCTGTTGTGA
GGATTAAATGAGATGATTCGTCTGAAGTGAATAAATCGTGTCTGGCACTGA

SEC. ID nº: 24

7127-7279

GGTGACTCTGGGGACCCCTGGTCTGTGGGGCATCCTGCAGGGCATTGTGTCTGGGGTGACGTCCCTTGTGACA
ACACCACCAAGCCTGGTGTCTATACCAAAGTCTGCCACTACTTGGAGTGGATCAGGAAACCATGAAGAGGAAC TG
A

SEC. ID nº: 25

Zyma

MKKLMVVL SLIAAAWAEQNKLVHGGPCDKTSHPYQAALYTSQHLLCGGVLIHPLWVLTAAHCKKPNLQVPLGKH
LRQRESSQEQSSVVRVAVIHPDYDAASHDQDIMLLRLARPAKLSELIQPLPLERDCSANTTSCHILGWGKTADGDFP
DTIQCAIYHLVSREECEHAYPGQITQNMLCAGDEKYGKDSQCQDGGPLVCGDHLRGLVSWGNI PCGSKEKPGVYT
NVCRTYTNWIKTIQAK

SEC. ID nº: 26

KLK-L4

MWPLALVIASLTALSGGVSQESSKVLNTNGTSGFLPGGYTCFPHSQPWQAALLVQGRLLCGGVLVHPKWVLTAAH
CLKEGLKVLGKHALGRVEAGEQREVVHSIPHEPYRRSPHLNHDHDI MLLELQSPVQLTGYIQTLP LSHNNRLT
PGTTTCRVSGWGTTS PQVNYPKTLQCANIQLRSDDECRQVYPGKITDNMLCAGTKEGGKDSCEGDSGGPLVCNRTL
YGI VSWGDFPCGQPD RGVYTRVSRVYLWIRETIRKYETQQQKWLKGPQ

SEC. ID nº: 27

KLK-L6

MFLLLTALQVLAIAMTQSQEDENKIIGHCTCRSSQPWQAALLA
GPRRRFLCGALLSGQWVITAAHCGRPIQLVALGKHNLRREATQQVLRVVRQVTHPNYNSRTHDNDMLLQLQPP
ARIGRAVRPIEVTQACASPGTSCRVSOGWTISSPIARYPASLQCVNINISPDEVCKAYPRITPGMVCAGVPQGG
KDSCQDGGSGGPLVCRGQLQGLVSWGMERCALPGYPGVYTNLCKYRSWIEETMRDK

ES 2 266 245 T3

SEC. ID n°: 28

TLSP

MQRLRWLRDWKSSGRGLTAAKEPGARSSPLQAMRILQLILLALATGLVGGETRIIKGFECCKPHSQPWQAALFEKTR
LLCGATLIAPRWLLTAAHCLKPRYIVHLGQHNLQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPV
SITWAVRPLTLSSRCVTAGTSCLISGWGSTSSPQLRLPHTLRANITIIHQKCEENAYPGNITDTMVCASVQEGGK
DSCQGDGGPLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYVDWIE

SEC. ID n°: 29

KLK-L3

MKLGLLALLSLLAGHWADTRAIGAECEPNSQPWQAGLPHLTRFCGATLISDRWLLTAAHCRKPYLWVRLGEH
HLWKWEGPEQLFRVTDFFPHPGFNKDLSDHNDIMLIRLPRQARLSPAVQPLNLSQTCVSPGMQCLISGWGAVS
SPKALFPVTLQCANISILENKLCHWAYPGHISDSMLCAGLWEGGRGSCQGDGGPLVCNGTLAGVVSOGAEPCSRP
RRPAVYTSVCHYLDWIEIMEN

SEC. ID n°: 30

NES1

MRAPHLHLSAASGARALAKLLPLLMAQLWAAEAALLPQNDTRLDPEAYGAPCARG SQPWQVSLFNGLSFH
CAGVLVDQSWVLTAACHGNKPLWARVGDDH LL-LLQG-EQLRRTT RSVVHPKYHQSGPI LPRRTDEHDLML
LKLARPVV-PGPRVR ALQLPYR-CAQPGDQ CQVAGWGTAAARRVK YNKGLTCSSITILSP
KECEVFYFGVVTNNM ICAGLDR-GQDPCQS DSGGPLVCDDETQGI LSWG-
VYPCGSAQHFAVYTQICKYMSWINK VIRSN

SEC. ID n°: 31

KLK-I.5

MGLSIFLLI.CVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLTAACHSCSRYWVRLGEHS
LSQLDWTEQIRIHSFGSVTHPGYLGASTSHEHDLRI.LRLRLPVRVTSSVQPLPLPNDCATAGTECHVSGWGTINHPR
NPPDLLQCLNI.SIVSHATCHGVYPGRIT'SNMVCAGGVPGQDACQGDGGPLVCGGVLQGI.VSWGVSVPCCQDGI

GVYTYICKYVDWIRMIMRNN

SEC. ID n°: 32

Neuropsina

MGRPRPRAAKTWMFLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGGQLLCGGVLVGGNWVLTAACHCKK
PKYTVRLGDHSLQNKDGPQEIPVVQSIHPFCYNSSDVEDHNDLMLLQLRDQASLGSKVKPISLADHCTQPGQKC
TVSGWGTVTSRENFPDTLNCAEVKIFPQKKCEDAYPGQITDGMVCAGSSKGADTCQGDGGPLVCDGALQGITSW
GSDPCGRSDKPGVYTNICRYLDWIKKIIGSKG

SEC. ID n°: 33

PSA

MWVPVVFVLTLSVTWIGAAPLILSRIVGGWECEKHSQPWQVLVASRGAHCGGVLVHPQWVLTAACHIRNKSVILLG
RHSLFHPEDTGQVQVSHSFPHPLYNMSLLKHQSLRPDESSHDMLLRLSEPAELTDAVKVMDLPTQEPALGTTTC
YASGWGSIEPEEFLTPKKLQCVDLHVISNDVCAQVHPQKVTKFMLCAGRWTGGKSTCSGDSGGPLVCNGVLQGITS
WGSEPCALPERPSLYTKVVHYRKWIKDTIVANP

SEC. ID n°: 34

HK2

MWDLVLSIALSVGCTGAVPLIQSRIVGGWECEKHSQPWQVAVYSHGWAHCGGVLVHPQWVLTAACHLKKNSQVWL
RHNLFEPEDTGQVQVSHSFPHPLYNMSLLKHQSLRPDESSHDMLLRLSEPAKITDVVKVLGLPTQEPALGTTTC
YASGWGSIEPEEFLRPRSLQCVSLHLLSNDMCARAYSEKVTDFMLCAGLWTGGKDTCCGGDSGGPLVCNGVLQGITS
WGPEPCALPEKPAVYTKVVHYRKWIKDTIAANP

SEC. ID n°: 35

HK1

MWFLVLCLALSIGGTGAAPPIQSRIVGGWECEKHSQPWQAALYHFSTFCGGILVHRQWVLTAACHISDNQYLWL
RHNLFDDENTAQFVHVSEFPHPGFNMSLLENHTRQADEYSHDLMLLRLTEPADTITDAVKVVELPTEPEVGS
CLASGWGSIEPENFSFPDDLQCVDLKILPNDECKKAHVQKVTDFMLCVGHLEGGKDTCVGDSGGPLMCDGVLQGV
SWGYPVCGTTPNKPSVAVRVLVSVKWIETIAENS

ES 2 266 245 T3

SEC. ID n°: 36

KLK-L2

MATARPPWMWVLCALITALLGVTEHVLANNVSCDHPSNTVPSGSNQDLGAGAGEDARSDDSSSRIINGSDCDMH
TQPWQAALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVRLGHYSLSPVYESGQMFQGVKSI PHPGYSHPGHSN
DLMLIKLNRRIPTKDVVPINVSSHCP SAGTKCLVSGWGTTKSPQVHF PKVLQCLNISVLSQKRCE DAYPRQIDDT
MFCAGDKAGR DSCQGD SGGPVVCNGSLQGLVSWGDYPCARP NRPGVYTNLCKFTKWIQETIQANS

SEC. ID n°: 37

Prostasa

MATAGNPWGWF LGYLILGVAGSLVSGSCSIINGEDCSPHSQPWQAALVMENELFCSGVLVHPQWVLSAAHCFQNS
YTIGLGLHLEADQEPGSQMV EASLSVRHPEYNRP LLANDLMLIKLDESVS SDTIR SISI ASQCPTAGNSCLVSGW
G LLANGRMPTVLQCVNVSVVSEEVCSKLYDPLYHPSMFCAGGGHDQK DSCNGD SGGPLICNGYI.QGLVSFGKAPCG
QGVPGVYTNLCKFTTEWIEKTVQAS

SEC. ID n°: 38

HSCCE

MARSLLLPLQILLLSLALETAGEEAQGDKIIDGAPCARGSHPWQVALLSGNQLHCHSCCEGGVLVNERWVLTAAHC
KMNEYTVHLGSDTLGDRAQRIKASKSFRHPGYSTQTHVNDLMLVKLNSQARLSSMVKKVRLPSRCEPPGTCTVS
GWGTTTSPDVTFPDL MCV DVKLISPQDCTKVYKD LLENSMLCAGIPDSKKNACNGD SGGPLVCRGTLQGLVS
WGTFPCGQPNDPGVYTQVCKFTKWINDTMKKHR

SEC ID n°: 39

CACAACGAGCCTGGGACCGCTGGG

SEC ID n°: 40

ATTAAA

SEC ID n°: 41

Tabla 1 KLK1-A

ATCCCTCCATTCCCATCTTT

SEC ID n°: 42

Tabla 1 KLK1-B

CACATACAATTCTCTGGTTC

SEC ID n°: 43

Tabla 1 KLK2-A

AGTGACACTGTCTCAGAATT

SEC ID n°: 44

Tabla 1 KLK2-B

CCCCAATCTCACGAGTGCAC

SEC ID n°: 45

Tabla 1 E5-A

GTCGGCTCTGGAGACATTC

SEC ID n°: 46

Tabla 1 E5-B

ES 2 266 245 T3

AACTGGGGAGGCTTGAGTC

SEC ID nº: 47

Tabla 1 KLK15-F1

CTCCTTCCTGCTGOCATCCA

SEC ID nº: 48

Tabla 1 KLK15-R1

ATCACACGGGTGGTCATGTG

SEC ID nº: 49

Tabla 1 KLK15-F2

CAAGTGGCTCTCTACGAGCG

SEC ID nº: 50

Tabla 1 KLK15-R2

GACACCAGGCTTGGTGGTGT

ES 2 266 245 T3

LISTA DE SECUENCIAS

5 <110> Mount Sinai Hospital
 Yousef, George, M.
 Diamandis, Eleftherios, P.

<120> Nuevo gen de la calicreína

10 <130> 3153-256
 <140> PCT/CA01/01141
 <141> 2001-08-10

15 <150> US 60/224,853
 <151> 2000-08-11

<160> 50

20 <170> Patente en version 3.1

<210> 1
25 <211> 8735
 <212> ADN

<213> *Homo sapiens*
30

35

40

45

50

55

60

65

ES 2 266 245 T3

<400> 1

	agaatgggtg ctgtgggatt caggggagac acctgttagg tgttggggcc tcccagaaga	60
5	ggtgggggca gagtgtcaga ggacaaagat gaatttgaa gatatgggga agaaggattt	120
	caattcacc ctaaagcttc ctgaggcctc ccgtgggtcg ggccctgcag tactggagac	180
	ccagagtga gtcagaccag ctctcgggg agctgccagt ctctagggg aggcagacac	240
10	cactgagggt caggggaggt cagagaaggc ctcaaggagg aagcggggct ggaagggaat	300
	ggcgttgat atgcggtggg aggaatagcc taagcatgaa atggcaggag ggaaaatggc	360
	agcactggct gcgtctagga caaggtcatg ggagaccag ggagaggggc tggaaggaa	420
15	gaagccactt ttgtccttga aagtgaggct ggagccaggc aactcatgcc tgtaatccca	480
	gcactttggg aggctgaggc ggttggtatca ctgaggtca ggagttcaag accagcctgg	540
	ccaacatggt gaaactccgt ctctactaaa attacaaaaa ttagctgggc gtggtggcac	600
20	acacctgtaa tcccgaattgc ttgggaggct gaggcaggag aatctcttga acccagaagg	660
	cagaggttac agtgagcga gatcacgcca ctccactcca acctgggcta cagagccaga	720
25	ctccgtctca aaaaaaaaa aaaaaaagaa aaaaaaagaa agaaagtga tttgaagagc	780
	tggactttat cctgggtggtg ccaaggatcc atggagggtg gtgagcaggg gaggggcaca	840
	gccagctcca gatgtagaaa gaccttttgg ggtcatggct ggagggcaag ctggtggagg	900
30	ggactggact ggagggggac caaaaggcc agataaggg gttgagatag accaggcgcg	960
	gtggctcatg cctgtaatcc cagcactttg ggaggccgag gtgggtggat catgaagtca	1020
	agagattgag gccatcctgg ctaaacacggt gaaacctgt ctctacttaa aaaaaaaaaa	1080
35	tttccaaaaa attagccggg cacggtggtg ggcgcctgta gtcccagcta ctcgagggc	1140
	tgaggcggga gaatggtgtg aacctgggag gtggagcttg cagtgaaccg acattgtgcc	1200
	actgcactcc agcctgggtg acagagttag actcctctc aaaaaataa aaaaagtgg	1260
40	gacagggggt ccttgcgtga tgatggagag agatccacc gctggtagca tgggtgctga	1320
	ggctgacagg tggaggagggt ggggcagggt ctgtccgagt gcctagagga agagtaaacc	1380
	ttccagagat gggggaccca gaaggaagcg cagagtgggg ttgggggaag gggataccgg	1440
	tggtcagaag aaatttatta acagtggatg ggataagtct gtgtctggag ggatcctggt	1500
45	ggaggcagaa gggctcctgcc tcacctggat tctctcactc cctccccaga ctgcagccga	1560
	acctgggtcc ctctccaca atgtggcttc tctcactct ctcttctctg ctggcatcca	1620
	caggtagggt ggccccagga gggggccagg tctgtgggag caggtgcccc ctccccagc	1680
50	atgtctgggc ccagtgatct gccagccctt acctaccca gagaccacta aagatccttc	1740
	cttcaccctc cacctgtgcc aatgtcccta agcccttacc gtcagggtgct ggtgctgctg	1800
	ctctggagtc gctatgttgc ctggggcctc tcgctgcca cgacaaggaa cacggtcctg	1860
55	gggttacaca aacctgagct gagtctggg gcaaccgctt ccttgcgtg tgtccttgag	1920
	ggaactgctt cacctctctg ggcttcgaat gccttctcta taagacagca cccacttgag	1980
	acaataacag tgagggtctca atagcataac agaggtaata tacatagcaa gcattagaca	2040
60	agtgcctgaga ggccaacagc acagacagac tccagcttga gtccacacc tgccactccc	2100
	tgtctcttac agggctcttg aggggattaa atgtggttgt gtgtgaggca gaagcataag	2160
	cctggcccag gtagtgcccc ttcagggtgtg caagccaggc acgggtgctta gagcttacat	2220

65

ES 2 266 245 T3

	acaacgtcta tgtgtggtgg gcaccaccga cctcatttga caaggggaagg ggctgtggct	2280
	cagaggggacg gccacaacat caagggtcacc ttgggtgtca ggcaaactcc agattgaact	2340
5	cagctgccac acaccaagaa attaatgtta acctgatgcc tctcttcttg agaaattggg	2400
	gggtggactt tcattaacgt tctgccacaa atgacctca ctctggggg cccctgagac	2460
	ccccacgcct ccagcctccc ctccggctct ctctgtgcac tcacctacct gcctcgcgcc	2520
10	tgcctgtgc gccagctgg gccctccacc ttctcttggc ttggactggc caggtgcagc	2580
	ctcgggtccc agctgttcag cccgtacct ccgcccttcg gaggacgacc tcaccttcc	2640
	tttgtaagc cccttgtcca ccacatccgc attcccttgg tctcacgggg gcctttggcc	2700
15	cagttcctga ctgtgatggg gagagtgtgg gcatttggtc tggtgtgca aatcctgccc	2760
	ctgtgtgggt gggagtgtgc atggcttcaa ccttcagggg atgcatccac attgccagt	2820
	ggagaggggt cctggtcctg tgacctgaa tgtctctaata catgtcctta agcataatgc	2880
20	cattctgtgt gtgtgtgtgt gtgtgtgtgt gtacatgcac gtgtgcagtg ggtatacaag	2940
	gccctgtatg ttcacatcct ctccacatgc atgagccaga tccccatatg tgaaacccaa	3000
	tcagtgactc cacagatctg gcttgggggc tgatctagag atggataaat atgtcctgcc	3060
25	ctggctgcct ctggcttcag ctgcatgtct ttgacctga atgccagcc ccgtgtcttg	3120
	gtgtgcccc agacagcaag tccacatctg agtggtggcc ttctgggttg gtgtctgcag	3180
	ctctaactct acaaaatgtc ttgtgggtga atcacggttt taaccttgac ttttttttgt	3240
30	ttgttttggt ttttttgaga cggagtctcg ctctgccgcc caagctggag ttcagtgggtg	3300
	caacctcagc tcactgcaac ctccgcctcc caggttcaag caattctgtc tctgcctccc	3360
	gagtagctag aattacaggc acgcaccacc acgccagct gatTTTTTgta tttttattta	3420
35	tttattttatt ttttttttag tagagacggg atttcacgat gttggccagg ctggtctcaa	3480
	actcctgacc tcaggtgate caccacctc ggcccttggc tcccaaagtg ctgggattac	3540
	aggcgtgagc caccacacct ggccaacctt gactatttat tataggtaat tctgtgcaga	3600
40	tgcttgactt atgttggcca tctccaggat ggacctgaac ttccacacgt atgtccctgt	3660
	gactaaatcc aggtgtcatt tgcaaaaaac aactaatatt attaatagc taccagggct	3720
	aggtatcact caccatacat acacacatgc acacacacac atacacattc ctacctcatc	3780
45	cttacaacaa tcttcatttt acagatgagg aaacagaggc acagacaggt cgaataactt	3840
	actcaaagtt tcacagctag tacattcgaa ccaggctta aggaccatc tttgtccaga	3900
	ccctgtatgc aagtgtctgt gacactggat gccaaagact acactagaga tgttgaattt	3960
50	aggtctgaac aatatccaat tctgtgtgtg tgtttgtgtg tgcatgtgtg tgtgtgtatg	4020
	tattcatgtc ttaacctacc atattcatac acacatatga acatctgtgc tgtgattctt	4080
	tttttttttt tttttttttt tttagatgg agtttctctc ttgtcaccca ggctggagtg	4140
	caatggagca acctccgctc actgcaatct ccgcctcccg ggttcaagcg attttcttgc	4200
55	ctcagcctcc agagtagctg ggattacagg caccgccac catgccagc taattttttg	4260
	tatttttgtt agagacaggg ttccccata ttggccaggc tggctctgaa ctctgacct	4320
60		
65		

ES 2 266 245 T3

	cagggtgatcc acccgccctcg gcctcccaaa gtgctgggat tacaggcatg agccaccgtg	4380
	cccagcctgt gctgtgatcc ttgaagctgc aaccatgtg catgcaagtg aatttcagct	4440
5	tccagtcctg tccatagctg tacctaagtg tggaaagctgg atgtgcatgt atgcatgtcc	4500
	atgaccttgt atagccacat ctgggactca tactgcacac tgaatttggc tgacatgtcc	4560
	agactctggg gccaaaggctg ggtcacacat actgagtggc cacatgcgtt tgacgtctgt	4620
10	gacaatttgg tgaccgtgaa tgactggttt caagtgacca cctgtctgaa cctgtatcca	4680
	gtgcccctgt ctccaccccc aaccacagag gacttcttgc cctctggtct gttccccttc	4740
	ctctctctcc cagagtctta tagcaaatgg ggtgggggct agagtcttgg agaaaacagg	4800
15	cagcggttgt aaataaacia cagggcaggc ggagcatggt ggctcacacc tgtaatccca	4860
	gcacttttgg aggtgagggc gggcagagca tttgaagtca gaagtttgag actacctggc	4920
	taacatggtg agacctcgtc tctactaaaa atacaaaaat tagccagggtg tgggtggcggg	4980
20	cacctcagct actcgggagg ctgaggcagg aggatcactt gaaccaggga ggcggaagtt	5040
	gcagtgagct gagatcatgc cactgcactc cagcctgggc aaaagagtga gactccgtct	5100
	caaaaacaac aacaacaaca aaacaaaaaa cagggcaggg tgtcttgaga agttagggga	5160
25	aaggcatagg catatagtag ttagggcagg gtgcaaggaa ggtgtaggag gcaatgtaa	5220
	cgccccctgt ctcaggcatc ctctaccctt tctcttagca gccaggatg gtgacaagtt	5280
	gctggaaggt gacgagtgtg caccctactc ccagccatgg caagtggctc tctacgagcg	5340
	tggacgcttt aactgtggcg cttccctcat cttccacac tgggtgctgt ctgcggccca	5400
30	ctgccaaagc cggatgaag gcaggggctc aggtctctga gggagcctgg ttcgggggga	5460
	agagctccta gatttggggg aagacggagg cagacgccag aactcctggg tctgaaaga	5520
	cgaggaggcc ggatgtcaag cccctgggtt aggaaggagt gtgtgtttca aagccttcga	5580
35	tctctgaagg aggaaggaga agactagtcc cagcttttga gcctcagttc tagggatgtg	5640
	agaatcctgg attcggggac agaccaggag ggggctggga gtagttggag gggatcagat	5700
	tctaggagtg tgctgactt cagactcgtt ggtccttag gagcaggggc tggaaaccatt	5760
40	ggcttcaggg tcttgggaaa aggtaatggg atgtcgagat ttctaaaggg tcggggagacc	5820
	tcgggttgcc cactctttga tctttctgtc ctctacttgc gggtaaccac tggcccgcc	5880
	tcactggcg ggaaaaccac tcgcccgcac agcttcatga gagtgcgcct gggagagcac	5940
45	aacctgcga agcgcgatg cccagagcaa ctacggacca cgtctcgggt cattccacac	6000
	ccgcgctacg aagcgcgcag ccaccgcaac gacatcatgt tgctgcgcct agtcagccc	6060
	gcacgcctga acccccaggc gcgccccgcg gtgctacca cgcgttgccc ccaccgggg	6120
50	gaggcctgtg tgggtgtctg ctggggcctg gtgtccaca acgagcctgg gaccgctggg	6180
	agccccgggt cacaaggtgc gtgaaaggat ggagctggat gcgaggcctc aaggaatcct	6240
	atgtctcagg gctcttgggc ggaggggaca agggccggaa tttatggatc tgctccaagt	6300
55	ccactgtctt cccagtgag tctcccagat acgttgcatg gtgccaacat cagcattatc	6360
	tcggacacat cttgtgaca gagctacca gggcgctga caaacacat ggtgtgtgca	6420
	ggcgcggagg gcagaggcgc agaactcgtg gaggtcagag cctagagggg ccatcaggcg	6480
60	gaagaagagg gatggggaca ggtgtgggag tccggatggg gttggatttt ctttgctttg	6540
	ggccagagaa gatgctaggg ttaggcttgg agatggagta ggaagagaag ttagaatagg	6600

ES 2 266 245 T3

ggtgaggttg gagttggggg tataggtggg gattgcgttg tttgaggtgg ataactgtga 6660
 tagttagttt gagatggcat ggggtggggg tgagaatggg aatggtttg tttgattctg 6720
 5 ggtgggaaat acgtcagggg tgaattggga tgaggtagat tttgtttgga atgcagaaga 6780
 catgaagatt gagattggat tttgagatgg gcatgggttt gatttgattt tgaatgggta 6840
 ggatgtgggc tgagttggat ttaacttagt acagttgcac tggagttgca tgggggtgag 6900
 10 attggatata ggttgggtga gttgtattga gctgtgttga attgggggtt ggggtggggg 6960
 tgggttggtt ctgtttggga taaactgggc tgtattgagt tgagttgggt tgggggtccc 7020
 tgggatgggg atggattggg tttgggggta gattgcaa atgtgattagg atgaggatga 7080
 15 atccaggagg tttcactcaa cctgagaccc cctcttttcc ccacaggggtg actctggggg 7140
 acccctgggc tgtgggggca tcttcagagg cattgtgtcc tgggggtgacg tcccttgtga 7200
 caacaccacc aagcctgggtg tctataccaa agtctgccac tacttggagt ggatcagggg 7260
 aaccatgaag aggaactgac tattctagcc tatctcctgt gcccctgact gagcagaagc 7320
 20 cccacagct ggccagcagc ccgcctgac atggaacaga acggagccat cccccaagac 7380
 cctgtccaag gccagatgt tagccaagga cttgtccac ctgaggacaa agctggcgct 7440
 caaggtcacc tgtttaatgc caagataaca aagcgtgat ccaagttgct ctgtaggaa 7500
 25 ttctgtgact tttttctggg gtcaaagaga aaccccgaga cactgtacac tgttctttt 7560
 caccacacc cccgatccct aggtgaggag aagcggcttg aagcagggct ccattcatc 7620
 aacacacatg accacccgtg tgatcttgaa caagaggccc aatctcactt cgccttgggt 7680
 30 tcccttatctg taaaatgaga ccatcttatt gctgacttca aagggtgtgt gtgaggatta 7740
 aatgagatga ttctgtctga ctgattaaaa tctgtgtctg cactgagtaa ataccctcta 7800
 tctctggatc ccagttaaag gacctaacag acactagatt accaagaatg gctttttctt 7860
 35 taaggtttag ttctgggccc ggcattgggtg ctcacacctg taatcccagc actttgggag 7920
 gccaaaggcg gcggtcact tgaggtcagg agtgcaagac cagcctggcc aacatgggtga 7980
 aaccccatct ctactaaaaa tactaaaaa atttagccgg gcgtgggtgg acacgactgt 8040
 40 aatcctagct acttgggagg gtgatgtggg aggatcgctt gaacttagga ggcaggagt 8100
 gcagtgagcc gagatcgcg cactgcactc cagcctgggtg acagagcaag actccatctc 8160
 agaaaaaaaa aaaaaaaaaa aaagatttag ttctgggctt cctggtagcc atggcaaaaa 8220
 45 ggcaaaact gtccttttct tagccaggtc cctgatatac agcagaggct ggaactctga 8280
 gctgctttga ttttaccaaa aagccaagac aacctgttg aagcctatgg gtttaccatt 8340
 gaggtgcag gaatctagtt cctaattatc ttcagagacc aaaaatgtg atgttcaagg 8400
 50 tctgtgaatg ttgaagtaca tgaacctggc tctgagacc taaatattgt actgggtgtg 8460
 ggggggaagg gtcattggaa tctgtgttga gcctgatctt gacctgcgag ggaaggttgt 8520
 ccagatctct ggactttgga ggaccgacgt tgagcaccat aatgggagca gaagtgcgag 8580
 55 gtctttgaga cccgcttgt tggggcgggc ccggatttg atgctaaaaa ttacctggga 8640
 accctgaata catctgggtt gggcgacaaa tgtgtggctc cccacacatc tttaggaaca 8700
 catttgggca acccggtggg agtgaacggc ctggc 8735

60 <210> 2
 <211> 1278
 <212> ADN
 <213> *Homo sapiens*
 65

ES 2 266 245 T3

<400> 2

	atgtggcttc tectactct ctccttctg ctggcatcca cagcagccca ggatggtgac	60
5	aagttgctgg aaggtgacga gtgtgcaccc cactcccagc catggcaagt ggctctctac	120
	gagcgtggac gctttaactg tggcgcttcc ctcactctcc cactactgggt gctgtctgag	180
	gcccactgcc aaagccgctt catgagagtg cgcctgggag agcacaacct gcgcaagcgc	240
10	gatggcccag agcaactacg gaccacgtct cgggtcattc cacacccgcg ctacgaagcg	300
	cgcagccacc gcaacgacat catgttgctg cgcctagtcc agcccgacg cctgaacccc	360
	caggtgcgcc ccgcggtgct acccagcgtg tgccccacc cgggggaggc ctgtgtggtg	420
15	tctggctggg gcctgggtgc ccacaacgag cctgggaccg ctgggagccc ccggtcacia	480
	gtgagtctcc cagatacgtt gcattgtgcc aacatcagca ttatctcgga cacatcttgt	540
	gacaagagct acccagggcg cctgacaaac accatgggtg gtgcaggcgc ggagggcaga	600
	ggcgcagaat cctgtgaggg tgactctggg ggacccctgg tctgtggggg catcctgcag	660
20	ggcatttgtt cctgggggtg cgtcccttgt gacaacacca ccaagcctgg tgtctatacc	720
	aaagtctgcc actacttgga gtggatcagg gaaaccatga agaggaactg actattctag	780
	cctatctcct gtgcccctga ctgagcagaa gccccacag ctggccagca gccccgctg	840
25	acatggaaca gaacggagcc atcccccaag accctgtcca agggccagat gttagccaag	900
	gacttgctcc acctgaggac aaagctggcg ctcaaggcca cctgtttaat gccaagataa	960
	caaagcgctg atccaagttg ctctgttaga atttctgtga cttttttctg gggtaaaga	1020
30	gaaacccga gacactgtac actgttccct ttacccacc accccgatcc ctagggtgag	1080
	agaagcggtt tgaagcaggg ctccattcat tcaacacaca tgaccaccgc tgtgatcttg	1140
	aacaagaggg ccaatctcac ttgccttggt ttcccttacc tgtaaaatga gaccatctta	1200
35	ttgctgactt caaagggctg ttgtgaggat taaatgagat gattcgtctg aactgattaa	1260
	aatcgtgtct ggcactga	1278

40 <210> 3

<211> 1124

<212> ADN

45 <213> *Homo sapiens*

<400> 3

	atgtggcttc tectactct ctccttctg ctggcatcca cagcagccca ggatggtgac	60
50	aagttgctgg aaggtgacga gtgtgcaccc cactcccagc catggcaagt ggctctctac	120
	gagcgtggac gctttaactg tggcgcttcc ctcactctcc cactactgggt gctgtctgag	180
	gcccactgcc aaagccgctt catgagagtg cgcctgggag agcacaacct gcgcaagcgc	240
55	gatggcccag agcaactacg gaccacgtct cgggtcattc cacacccgcg ctacgaagcg	300
	cgcagccacc gcaacgacat catgttgctg cgcctagtcc agcccgacg cctgaacccc	360
	cagtgagtct ccagatacgt ttgcattgtg ccaacatcag cattatctcg gacacatctt	420

60

65

ES 2 266 245 T3

5 gtgacaagag ctacccaggg cgcctgacaa acaccatggt gtgtgcaggc gcggagggca 480
 gaggcgcaga atcctgtgag ggtgactctg ggggacccct ggtctgtggg ggcatcctgc 540
 agggcattgt gtcctggggg gacgtccctt gtgacaacac caccaagcct ggtgtctata 600
 ccaaagtctg ccactacttg gagtggatca gggaaacccat gaagagggaac tgactattct 660
 agcctatctc ctgtgcccct gactgagcag aagccccac agctggccag cagccccgcc 720
 10 tgacatggaa cagaacggag ccatcccca agaccctgtc caaggcccag atgttagcca 780
 aggacttgtc ccacctgagg acaaagctgg cgtcaagggt cacctgttta atgccaagat 840
 aacaaagcgc tgateccaagt tgctctgtag gaatttctgt gacttttttc tggggtcaaa 900
 gagaaacccc gagacactgt acactgttcc ttttcaccca ccccccgat ccctaggtga 960
 15 ggagaagcgg cttgaagcag ggctccattc attcaacaca catgaccacc cgtgtgatct 1020
 tgaacaagag gccaatctc acttcgcctt ggtttcctta tctgtaaaat gagaccatct 1080
 tattgctgac ttcaaagggc tgttgtgagg attaaatgag atga 1124

<210> 4

<211> 1105

<212> ADN

25 <213> *Homo sapiens*

<400> 4

 atgtggcttc tcctcactct ctccttcttg ctggcatcca cagcagccca ggatggtgac 60
 30 aagttgctgg aagggtgacga gtgtgcaccc cactcccagc catggcaagt ggctctctac 120
 gagcgtggac gctttaactg tggcgcttcc ctcactctcc cactctgggt gctgtctgcy 180
 gcccactgcc aaagccgctt catgagagtg cgcctgggag agcacaacct gcgcaagcgc 240
 35 gatggcccag agcaactacg gaccacgtct cgggtcattc cacaccgcg ctacgaagcg 300
 cgcagccacc gcaacgacat catgttgctg cgcctagtcc agcccgacg cctgaacccc 360
 caggtgcgcc ccgcggtgct acccacgcgt tgccccacc cgggggaggc ctgtgtggtg 420
 40 tctggctggg gcctgggtgt ccacaacgag cctgggacgc ctgggagccc ccggtcacaa 480
 gggtgactct gggggacccc tgggtctgtg gggcatcctg cagggcattg tgcctggggg 540
 tgacgtccct tgtgacaaca ccaccaagcc tgggtgtctat accaaagtct gccactactt 600
 ggagtggatc agggaaacca tgaagaggaa ctgactattc tagcctatct cctgtgcccc 660
 45 tgactgagca gaagccccca cagctggcca gcagccccgc ctgacatgga acagaacgga 720
 gccatcccc aagacctgt ccaaggccca gatgttagcc aaggacttgt cccacctgag 780
 gacaaagctg gcgtcaagg tcacctgttt aatgccaaga taacaaagcg ctgatccaag 840
 50 ttgctctgta ggaatttctg tgactttttt ctgggggtcaa agagaaaccc cgagacactg 900
 tacactgttc cttttcaccc accacccgga tccttaggtg aggagaagcg gcttgaagca 960
 gggctccatt cattcaacac acatgaccac ccgtgtgac ttgaacaaga ggcccaatct 1020
 55 cacttcgcct tggtttcctt atctgtaaaa tgagaccatc ttattgctga cttcaaaggg 1080
 ctgttgtgag gattaaatga gatga 1105

<210> 5

60 <211> 987

<212> ADN

<213> *Homo sapiens*

65

ES 2 266 245 T3

<400> 5

```

5      atgtggcttc tcctcactct ctcccttctg ctggcatcca cagcagccca ggatgggtgac    60
      aagttgctgg aaggtgacga gtgtgcaccc cactcccagc catggcaagt ggctctctac    120
      gagcgtggac gctttaactg tggcgcttcc ctcatctccc cacactgggt gctgtctgcg    180
      gcccactgcc aaagccgctt catgagagtg cgcctgggag agcacaacct gcgcaagcgc    240
10     gatggcccag agcaactacg gaccacgtct cgggtcattc cacacccgcg ctacgaagcg    300
      cgcagccacc gcaacgacat catgttgctg cgcctagtcc agcccgcacg cctgaacccc    360
      cagggtgact ctgggggacc cctgggtctgt gggggcatcc tgcagggcat tgtgtcctgg    420
15     ggtgacgtcc cttgtgacaa caccaccaag cctgggtgtct ataccaaagt ctgccactac    480
      ttggagtgga tcagggaaac catgaagagg aactgactat tctagcctat ctccgtgtgcc    540
      cctgactgag cagaagcccc cacagctggc cagcagcccc gcctgacatg gaacagaacg    600
20     gagccatccc ccaagaccct gtccaaggcc cagatgttag ccaaggactt gtcccacctg    660
      aggacaaagc tggcgctcaa ggtcacctgt ttaatgccaa gataacaaag cgctgatcca    720
      agttgctctg taggaatttc tgtgactttt ttctggggtc aaagagaaac cccgagacac    780
25     tgtacactgt tccttttcac ccaccacccc gatccctagg tgaggagaag cggcttgaag    840
      cagggctcca ttcatccaac acacatgacc acccgtgtga tcttgaacaa gagggcccaat    900
      ctcaacttcg cttggtttcc ttatctgtaa aatgagacca tcttattgct gacttcaaag    960
30     ggctgttgtg aggattaaat gagatga                                     987

```

<210> 6

<211> 256

35 <212> PRT

<213> *Homo sapiens*

<400> 6

```

40     Met Trp Leu Leu Leu Thr Leu Ser Phe Leu Leu Ala Ser Thr Ala Ala
      1          5          10          15

45     Gln Asp Gly Asp Lys Leu Leu Glu Gly Asp Glu Cys Ala Pro His Ser
      20          25          30

      Gln Pro Trp Gln Val Ala Leu Tyr Glu Arg Gly Arg Phe Asn Cys Gly
      35          40          45

50     Ala Ser Leu Ile Ser Pro His Trp Val Leu Ser Ala Ala His Cys Gln
      50          55          60

55     Ser Arg Phe Met Arg Val Arg Leu Gly Glu His Asn Leu Arg Lys Arg
      65          70          75          80

```

60

65

ES 2 266 245 T3

Asp Gly Pro Glu Gln Leu Arg Thr Thr Ser Arg Val Ile Pro His Pro
 85 90 95
 5 Arg Tyr Glu Ala Arg Ser His Arg Asn Asp Ile Met Leu Leu Arg Leu
 100 105 110
 Val Gln Pro Ala Arg Leu Asn Pro Gln Val Arg Pro Ala Val Leu Pro
 115 120 125
 10 Thr Arg Cys Pro His Pro Gly Glu Ala Cys Val Val Ser Gly Trp Gly
 130 135 140
 15 Leu Val Ser His Asn Glu Pro Gly Thr Ala Gly Ser Pro Arg Ser Gln
 145 150 155 160
 Val Ser Leu Pro Asp Thr Leu His Cys Ala Asn Ile Ser Ile Ile Ser
 165 170 175
 20 Asp Thr Ser Cys Asp Lys Ser Tyr Pro Gly Arg Leu Thr Asn Thr Met
 180 185 190
 25 Val Cys Ala Gly Ala Glu Gly Arg Gly Ala Glu Ser Cys Glu Gly Asp
 195 200 205
 Ser Gly Gly Pro Leu Val Cys Gly Gly Ile Leu Gln Gly Ile Val Ser
 210 215 220
 30 Trp Gly Asp Val Pro Cys Asp Asn Thr Thr Lys Pro Gly Val Tyr Thr
 225 230 235 240
 35 Lys Val Cys His Tyr Leu Glu Trp Ile Arg Glu Thr Met Lys Arg Asn
 245 250 255
 <210> 7
 <211> 121
 40 <212> PRT
 <213> *Homo sapiens*
 <400> 7
 45 Met Trp Leu Leu Leu Thr Leu Ser Phe Leu Leu Ala Ser Thr Ala Ala
 1 5 10 15
 50 Gln Asp Gly Asp Lys Leu Leu Glu Gly Asp Glu Cys Ala Pro His Ser
 20 25 30
 Gln Pro Trp Gln Val Ala Leu Tyr Glu Arg Gly Arg Phe Asn Cys Gly
 35 40 45
 55 Ala Ser Leu Ile Ser Pro His Trp Val Leu Ser Ala Ala His Cys Gln
 50 55 60
 60 Ser Arg Phe Met Arg Val Arg Leu Gly Glu His Asn Leu Arg Lys Arg
 65 70 75 80

ES 2 266 245 T3

Asp Gly Pro Glu Gln Leu Arg Thr Thr Ser Arg Val Ile Pro His Pro
 85 90 95
 5 Arg Tyr Glu Ala Arg Ser His Arg Asn Asp Ile Met Leu Leu Arg Leu
 100 105 110
 Val Gln Pro Ala Arg Leu Asn Pro Gln
 115 120
 10
 <210> 8
 <211> 161
 <212> PRT
 15 <213> *Homo sapiens*
 <400> 8
 20 Met Trp Leu Leu Leu Thr Leu Ser Phe Leu Leu Ala Ser Thr Ala Ala
 1 5 10 15
 Gln Asp Gly Asp Lys Leu Leu Glu Gly Asp Glu Cys Ala Pro His Ser
 20 25 30
 25 Gln Pro Trp Gln Val Ala Leu Tyr Glu Arg Gly Arg Phe Asn Cys Gly
 35 40 45
 30 Ala Ser Leu Ile Ser Pro His Trp Val Leu Ser Ala Ala His Cys Gln
 50 55 60
 Ser Arg Phe Met Arg Val Arg Leu Gly Glu His Asn Leu Arg Lys Arg
 65 70 75 80
 35 Asp Gly Pro Glu Gln Leu Arg Thr Thr Ser Arg Val Ile Pro His Pro
 85 90 95
 40 Arg Tyr Glu Ala Arg Ser His Arg Asn Asp Ile Met Leu Leu Arg Leu
 100 105 110
 Val Gln Pro Ala Arg Leu Asn Pro Gln Val Arg Pro Ala Val Leu Pro
 115 120 125
 45 Thr Arg Cys Pro His Pro Gly Glu Ala Cys Val Val Ser Gly Trp Gly
 130 135 140
 50 Leu Val Ser His Asn Glu Pro Gly Thr Ala Gly Ser Pro Arg Ser Gln
 145 150 155 160
 Gly
 55
 <210> 9
 <211> 171
 <212> PRT
 60 <213> *Homo sapiens*
 65

ES 2 266 245 T3

<400> 9

5 Met Trp Leu Leu Leu Thr Leu Ser Phe Leu Leu Ala Ser Thr Ala Ala
1 5 10 15

Gln Asp Gly Asp Lys Leu Leu Glu Gly Asp Glu Cys Ala Pro His Ser
20 25 30

10 Gln Pro Trp Gln Val Ala Leu Tyr Glu Arg Gly Arg Phe Asn Cys Gly
35 40 45

Ala Ser Leu Ile Ser Pro His Trp Val Leu Ser Ala Ala His Cys Gln
50 55 60

15 Ser Arg Phe Met Arg Val Arg Leu Gly Glu His Asn Leu Arg Lys Arg
65 70 75 80

20 Asp Gly Pro Glu Gln Leu Arg Thr Thr Ser Arg Val Ile Pro His Pro
85 90 95

Arg Tyr Glu Ala Arg Ser His Arg Asn Asp Ile Met Leu Leu Arg Leu
100 105 110

25 Val Gln Pro Ala Arg Leu Asn Pro Gln Gly Asp Ser Gly Gly Pro Leu
115 120 125

30 Val Cys Gly Gly Ile Leu Gln Gly Ile Val Ser Trp Gly Asp Val Pro
130 135 140

35 Cys Asp Asn Thr Thr Lys Pro Gly Val Tyr Thr Lys Val Cys His Tyr
145 150 155 160

Leu Glu Trp Ile Arg Glu Thr Met Lys Arg Asn
165 170

40 <210> 10

<211> 8

<212> PRT

45 <213> Secuencia artificial

<220>

<223> segmentos péptidos/aminoácido procedentes de secuencia humana

<400> 10

50 His Asn Glu Pro Gly Thr Ala Gly
1 5

55 <210> 11

<211> 1580

<212> ADN

<213> *Homo sapiens*

60

65

ES 2 266 245 T3

<400> 11

5	agaatgggtg ctgtgggatt caggggagac acctgttagg tgttggggcc tcccagaaga	60
	ggtaggggca gagtgtcaga ggacaaagat gaatttggaa gatatgggga agaaggattt	120
	caattcaccc tcaaagcttc ctgaggcctc ccgtgggtcg ggccctgcag tactggagac	180
	ccagagtggg gtcagaccag ctctcgggg agctgccagt ctctagggg aggcagacac	240
10	cactgagggt caggggaggt cagagaaggc ctcaaggagg aagcggggct ggaagggat	300
	ggcgttgat atgcggtggg aggaatagcc taagcatgaa atggcaggag ggaaaatggc	360
	agcactggct gcgtctagga caaggctcatg ggagaccag ggagaggggc tggaaaggaa	420
15	gaagccactt ttgtccttga aagtgaggct ggagccaggc aactcatgcc tgtaatccca	480
	gcactttggg aggtcaggc ggggtgatca ctagaggta ggaattcaag accagcctgg	540
20	ccaacatggt gaaactccgt ctctactaaa attacaaaaa ttagctgggc gtggtggcac	600
	acacctgtaa tcccaattgc ttgggaggct gaggcaggag aatctcttga acccagaagg	660
	cagaggttac agtgagcggg gatcacgcca ctccactcca acctgggcta cagagccaga	720
25	ctccgtctca aaaaaaaaaa aaaaaaagaa aaaaaaagaa agaaagtga tttgaagagc	780
	tggactttat cctgggtgtg ccaaggatcc atggaggggt gtgagcaggg gaggggcaca	840
	gccagctcca gatgtagaaa gaccctttgg ggtcatggct ggagggcaag ctggtggagg	900
30	ggactggact ggagggggac ccaaaaggcc agataagagg gttgagatag accaggcgcg	960
	gtggtcatg cctgtaatcc cagcactttg ggaggccgag gtgggtggat catgaagtca	1020
	agagattgag gccatcctgg ctaacacggt gaaaccctgt ctctacttaa aaaaaaaaaa	1080
35	tttccaaaaa attagccggg cacggtgtg ggcgcctgta gtcccagcta ctcgaggaggc	1140
	tgaggcggga gaatggtgtg aacctgggag gtggagcttg cagtgagccg acattgtgcc	1200
	actgcactcc agcctgggtg acagagttag actccgtctc aaaaaataa aaaaagttgg	1260
40	gacagggggt ccttgctgta tgatggagag agatccacc gctggttagca tgggtgctgga	1320
	ggctgacagg tggaggaggt ggggcagggt ctgtccgagt gcctagagga agagtaaacc	1380
	ttccagagat gggggaccca gaaggagcg cagagtggg ttgggggaag ggataaccgg	1440
	tggtcagaag aaatttatta acagtggatg ggataagtct gtgtctggag ggatcctgg	1500
45	ggaggcagaa gggtcctgcc tcacctggat tctctcactc cctccccaga ctgcagccga	1560
	accctgggtcc ctctccaca	1580

50 <210> 12

<211> 43

<212> ADN

55 <213> *Homo sapiens*

<400> 12

atgtgcttc tctcactct ctcttctctg ctggcatcca cag

43

60 <210> 13

<211> 3635

<212> ADN

65 <213> *Homo sapiens*

ES 2 266 245 T3

<400> 13

	gtgaggtggc cccaggaggg ggccaggtct gtgggagcag gtgccccctt cccaagcatg	60
5	tctgggcecca gtgatctgcc agccccctacc tcacccagag accactaaag atccttcctt	120
	caccctccac ctgtgccaat gtccctaagc ccttaccgtc aggtgctggt gctgctgctc	180
	tggagtcgct atgttgctcg gggcctctcg ctgcccacga caaggaacac ggtcctgggg	240
10	ttacacaaac ctgagctgag tcctggggca accgcttcct tgctgtgtgt ccttgaggga	300
	actgcttcac ctctctgggc ttccaatgcc ttctctataa gacagcacc ctttgagaca	360
15	ataacagtga ggtctcaata gcataacaga ggtaataatac atagcaagca ttagacaagt	420
	gctgagaggg caacagcaca gacagactcc agcttgagtc ccacacctgc cactccctgt	480
	ctcttacagg gtctttgagg ggattaaatg tggttgtgtg tgaggcagaa gcataagcct	540
	ggcccaggta gtgccccctc aggtgtgcaa gccaggcacg gtgcttagag cttacatata	600
20	acgtctatgt gtggtgggca ccaccgacct catttgacaa gggaaggggc tgtggctcag	660
	agggacggcc acaacatcaa ggtcaccttg ggtgtcaggc aaactccaga ttgaactcag	720
	ctgccacaca ccaagaaatt aattgtaacc tgatgcctct cttctggaga aattgggggg	780
25	tggactttca ttaacgttct gccacaaatg accctcactc ctggggggccc ctgagacccc	840
	cacgcctcca gcctccccct cggtctctctc tgtgactca cctacctgcc tcgcgcctgc	900
	ctgctgcgcc cagctggggc ctccaccttc ctctgcttg gactggccag gtgcagcctc	960
30	ggtgcccagc tgttcagccc gtacctccg ccttcggag gacgacctca ccttccttt	1020
	gttaagcccc ttgtccacca catccgcatt cccctggctc cacggggggc tttggcccag	1080
	ttcctgactg tgatggggag agtgtgggca tttggtctgg ctgtgcaa cctgcccctg	1140
35	tgtgggtggg agtgtgcatg gcttcaacct tcaggggatg catccacatt gccagtgga	1200
	gaggggtect ggtcctgtga ccttgaatgt ctctaatacat gtccttaagc ataatgccat	1260
	tctgtgtgtg tgtgtgtgtg tgtgtgtgt catgcacgtg tgcagtgggt atacaaggcc	1320
40	ctgtatgttc acatcctctc cacatgcatg agccagatcc ccatatgtga aacccaatca	1380
	gtgactccac agatctggct tgggggctga tctagagatg gataaatatg tcctgccctg	1440
	gctgcctctg gcttcagctg catgtctttg accttgatg cccagccccg tgtctgggtg	1500
45	ctgccccaga cagcaagtcc acatctgagt gttggccttc tgggttggtg tctgcagctc	1560
	taactctaca aaatgtcttg tgggtgaatc acgggtttta ccttgacttt ttttgtttg	1620
	tttggttttt tttgagacgg agtctcgctc tgccgccccaa gctggagtgc agtggtgcaa	1680
50	cctcagctca ctgcaacctc cgcctcccag gttcaagcaa ttctgtctct gcctcccag	1740
	tagctagaat tacaggcacg caccaccacg cccagctgat ttttgtattt ttattttatt	1800
	atttatttat ttttagtag agacgggatt tcacgatgtt ggccaggctg gtctcaact	1860
	cctgacctca ggtgatccac ccacctcggc cttggcctcc caaagtgtg ggattacagg	1920
55	cgtgagccac cacacctggc caaccttgac tatttattat aggtaattct gtgcagatgt	1980
	ctgacttatg ttggccatct ccaggatgga cctgaacttt cacacgtatg tccctgtgac	2040
	taaatccagg tgtcatttgc aaaaaacaac taatattatt aagtagctac cagggtcagg	2100
60	tatcactcac catacatata cacatgcaca cacacacata cacattccta cctcatcctt	2160
	acaacaatct tcattttaca gatgaggaaa cagaggcaca gacaggtcga ataacttact	2220

65

ES 2 266 245 T3

	caaagtttca cagctagtac attcgaaccc aggcttaagg acccatcttt gtccagaccc	2280
	tgtatgcaag tgtctgtgac actggatgcc aagactcaca ctagagatgt tgaatttagg	2340
5	tctgaacaat atccaattct gtgtgtgtgt ttgtgtgtgc atgtgtgtgt gtgtatgtat	2400
	tcatgtctta accatccata ttcataataca catatgaaca tctgtgctgt gattcttttt	2460
	tttttttttt tttttttttt gagatggagt ttactcttg tcaccaggc tggagtgcaa	2520
10	tggagcaacc tccgctcact gcaatctccg cctcccggt tcaagcgatt ttcctgcctc	2580
	agcctccaga gtagctggga ttacaggcac ccgccacccat gccagctaa tttttgtat	2640
	ttttgttaga gacaggggtt ccccatattg gccaggctgg tctcgaactc ctgacctcag	2700
15	gtgatccacc cgcctcggcc tcccaaagtg ctgggattac aggcatgagc caccgtgcc	2760
	agcctgtgct gtgattcttg aagctgcaac ccatgtgcat gcaagtgaat ttcagcttcc	2820
	agtcctgtcc atagctgtac ctaagtgtgg aagctggatg tgcattgtat catgtccatg	2880
20	accttgtata gccacatctg ggactcatac tgcacactga atttggtga catgtccaga	2940
	ctctggggcc aaggctgggt cacacatact gaggggccac atgcgtttga cgtctgtgac	3000
	aatttggtga ccgtgaatga ctgggtttcaa gtgaccacct gtctgaacct gtatccagt	3060
25	ccccgtctc caccaccaac cacagaggac ttcttgccct ctggtctgtt ccccttctc	3120
	tctctcccag agtcttatag caaatgggtt gggggctaga gttctggaga aaacaggcag	3180
	cgggtgtaaa taaacaacag ggcaggcgga gcatggtggc tcacacctgt aatcccagca	3240
30	ctttgggagg ctgaggcggg cagagcattt gaagtcagaa gtttgagact acctggctaa	3300
	catggtgaga cctcgtctct actaaaaata caaaaattag ccagggtgtg tggcgggcac	3360
	ctcagctact cgggaggctg aggcaggagg atcacttga cccaggaggc ggaagttgca	3420
35	gtgagctgag atcatgccac tgcactccag cctgggcaaa agagtgagac tccgtctcaa	3480
	aaacaacaac aacaacaaa caaaaacag ggcagggtgt cttgagaagt taggggaaag	3540
	gcataggcat atagtagtta gggcagggtg caagggaagg gtaggaggca atgtaaactg	3600
40	ccctgtctc aggcatctc tacccttct cttag	3635

<210> 14
 <211> 154
 <212> ADN
 <213> *Homo sapiens*
 <400> 14

50	cagcccagga tgggtgacaag ttgctggaag gtgacgagtg tgcacccac tcccagccat	60
	ggcaagtggc tctctacgag cgtggacgct ttaactgtgg cgcttccctc atctcccccac	120
55	actgggtgct gtctgcggcc cactgccaac gccg	154

<210> 15
 <211> 500
 <212> ADN
 <213> *Homo sapiens*

65

ES 2 266 245 T3

<400> 15

	gtatgaaggc aggggctcag ggtcctgagg gagcctgggt cggggggaag agctcctaga	60
5	tttgggggaa gacggaggca gacgccagaa ctctgggtt ctgaaagacg aggaggccgg	120
	atgtcaagcc cctgggttag gaaggagtgt gtgtttcaaa gccttcgac tctgaaggag	180
	gaaggagaag actagttcca gcttttgagc ctcaagttcta gggatgtgag aatcctggat	240
10	tcggggacag accaggaggg ggctgggagt agttggaggg gatcgagttc taggagtgtg	300
	cctgacttca gactcgttgg tccttgagga gcaggggctg gaaccattgg ctccagggtc	360
	ttgggaaaag gtaatgggat gtcgagattt ctaaagggtc gggagacctc gggttgcccc	420
15	ctctttgatc tttctgtcct ctacttgctg gtaaccactg gccgcactc cactggcggg	480
	aaaaccactc gccgcacag	500

<210> 16

<211> 284

20 <212> ADN

<213> *Homo sapiens*

<400> 16

25	cttcatgaga gtgcgcctgg gagagcaca cctgcgcaag cgcgatggcc cagagcaact	60
	acggaccacg tctcgggtca ttccacaccc gcgctacgaa gcgcgcagcc accgcaacga	120
	catcatgttg ctgcgcctag tccagcccg cgcctgaac cccaggtgc gcccgcgggt	180
30	gctaccacg cggtgcccc acccgggga ggctgtgtg gtgtctggct ggggcctggt	240
	gtcccacaac gagcctggga ccgctgggag ccccggtca caag	284

<210> 17

35 <211> 119

<212> ADN

<213> *Homo sapiens*

<400> 17

40	tgcgtgaaag gatggagctg gatgcgaggc ctcaaggaat cctatgctcc agggctcttg	60
	ggcggagggg acaagggccg gaatttatgg atctgtcca agtccactgt cttcccag	119

<210> 18

45 <211> 284

<212> ADN

<213> *Homo sapiens*

50 <400> 18

	cttcatgaga gtgcgcctgg gagagcaca cctgcgcaag cgcgatggcc cagagcaact	60
	acggaccacg tctcgggtca ttccacaccc gcgctacgaa gcgcgcagcc accgcaacga	120
55	catcatgttg ctgcgcctag tccagcccg cgcctgaac cccaggtgc gcccgcgggt	180
	gctaccacg cggtgcccc acccgggga ggctgtgtg gtgtctggct ggggcctggt	240
	gtcccacaac gagcctggga ccgctgggag ccccggtca caag	284

60 <210> 19

<211> 120

<212> ADN

65 <213> *Homo sapiens*

ES 2 266 245 T3

<400> 19		
	gtgcgtgaaa ggatggagct ggatgagagg cctcaaggaa tcctatgctc cagggctctt	60
5	gggcggaggg gacaagggcc ggaatttatg gatctgctcc aagtccactg tcttccccag	120
<210> 20		
<211> 137		
10	<212> ADN	
	<213> <i>Homo sapiens</i>	
<400> 20		
15	tgagtctccc agatacgttg cattgtgcc aacacagcat tatctcggac acatcttgtg	60
	acaagagcta cccagggcgc ctgacaaaca ccatgggtgtg tgcaggcgcg gagggcagag	120
	gcgcagaatc ctgtgag	137
20	<210> 21	
	<211> 673	
	<212> ADN	
25	<213> <i>Homo sapiens</i>	
<400> 21		
30	gtcagagcct agagggggcca tcaggcggaa gaagagggat ggggacaggt gtgggagtc	60
	ggatgggggt ggattttctt tgctttgggc cagagaagat gctaggggta ggcttgagga	120
	tggagtagga agagaagtta gaataggggt gaggttgag ttgggggttat aggtggggat	180
	tgcgttggtt gaggtggata actgtgatag ttagtttgag atggcatggg ttgggggtga	240
35	gaatgggaat ggtttggttt gattctgggt gggaaatacg tcagggttga attgggatga	300
	ggtagatctt gtttggaatg cagaagacat gaagattgag attggatttt gagatgggca	360
	tgggtttgat ttgattttga atgggtgagga tgtgggctga gttggattta acttagtaca	420
40	gttgcaactgg agttgcatgg ggggtgagatt ggatataggt tgggtgagtt gtattgagct	480
	gtgttgaatt ggggttgggg ttgggggttg gttggctctg tttgggataa actgggctgt	540
	attgagttga gttgggttg ggttccctgg gatggggatg gattgggttt ggggtgagat	600
45	tgcaaatggt gattaggatg aggatgaatc caggagggtt cactcaacct gagacccct	660
	cttttcccca cag	673
<210> 22		
50	<211> 1046	
	<212> ADN	
	<213> <i>Homo sapiens</i>	
55		
60		
65		

ES 2 266 245 T3

<400> 22

	tgcgccccgc ggtgctaccc acgcgttgcc cccacccggg ggaggcctgt gtggtgtctg	60
5	gctggggcct ggtgtccac aacgagcctg ggaccgctgg gagccccgg tcacaagggtg	120
	cgtgaaagga tggagctgga tgcgaggcct caaggaatcc tatgctccag ggctcttggg	180
	cggaggggac aagggccgga atttatggat ctgctccaag tccactgtct tccccagtga	240
10	gtctcccaga tacgttgcat tgtgccaaca tcagcattat ctcgacaca tcttgtgaca	300
	agagctaccc agggcgctg acaaacacca tgggtgtgtg aggcgaggag ggcagaggcg	360
	cagaatcctg tgaggtcaga gcctagaggg gccatcaggc ggaagaagag gatggggaca	420
15	ggtgtgggag tccggatggg gttggatttt ctttgctttg ggccagagaa gatgctaggg	480
	ttaggcttgg agatggagta ggaagagaag ttagaatagg ggtgagggtg gagttggggt	540
	tatagggtgg gattgcgttg tttgagggtg ataactgtga tagttagttt gagatggcat	600
20	gggttggggt tgagaatggg aatggttttg tttgattctg ggtgggaaat acgtcagggt	660
	tgaattggga tgaggtagat tttgtttgga atgcagaaga catgaagatt gagattggat	720
	tttgagatgg gcatgggttt gatttgattt tgaatgggtg ggatgtgggc tgagtggat	780
25	ttaacttagt acagttgcac tggagttgca tgggggtgag attggatata ggttgggtga	840
	gttgtattga gctgtgttga attgggggtg gggttggggt tgggttggct ctgtttggga	900
	taaactgggc tgtattgagt tgagtgggt tggggttccc tgggatgggg atggattggg	960
30	tttggggtga gattgcaaat ggtgattagg atgaggatga atccaggagg ttctactcaa	1020
	cctgagaccc cctcttttcc ccacag	1046

<210> 23

<211> 660

35 <212> ADN

<213> *Homo sapiens*

<400> 23

40	ggtgactctg ggggacccct ggtctgtggg ggcacccctg agggcattgt gtcctggggt	60
	gacgtccctt gtgacaacac caccaagcct ggtgtctata ccaaagtctg ccactacttg	120
	gagtgatca gggaaaccat gaagaggaaac tgactattct agcctatctc ctgtgcccc	180
45	gactgagcag aagccccac agctggccag cagccccgcc tgacatggaa cagaacggag	240
	ccatccccca agaccctgtc caaggcccag atgttagcca aggacttgtc ccacctgagg	300
	acaaagctgg cgctcaaggt cacctgttta atgccaagat aacaaagcgc tgatccaagt	360
50	tgctctgtag gaatttctgt gacttttttc tggggtcaaa gagaaacccc gagacactgt	420
	acactgttcc ttttcaccca ccaccccgat ccctagggtg ggagaagcgg cttgaagcag	480
	ggctccattc attcaacaca catgaccacc cgtgtgatct tgaacaagag gccaatctc	540
55	acttcgcctt gggttcctta tctgtaaaat gagaccatct tattgctgac ttcaaagggc	600
	tgtgtgagg attaaatgag atgattcgtc tgaactgatt aaaatcgtgt ctggcactga	660

<210> 24

60 <211> 153

<212> ADN

<213> *Homo sapiens*

65

ES 2 266 245 T3

<400> 24

```

          ggtgactctg ggggaccctt ggtctgtggg ggcatcctgc agggcattgt gtcttggggt      60
5      gacgtccctt gtgacaacac caccaagcct ggtgtctata ccaaagtctg ccactacttg      120
          gagtggatca gggaaacat gaagaggaac tga                                  153

```

<210> 25

10 <211> 244

<212> PRT

<213> *Homo sapiens*

<400> 25

```

15      Met Lys Lys Leu Met Val Val Leu Ser Leu Ile Ala Ala Ala Trp Ala
          1          5          10          15

20      Glu Glu Gln Asn Lys Leu Val His Gly Gly Pro Cys Asp Lys Thr Ser
          20          25          30

25      His Pro Tyr Gln Ala Ala Leu Tyr Thr Ser Gly His Leu Leu Cys Gly
          35          40          45

30      Gly Val Leu Ile His Pro Leu Trp Val Leu Thr Ala Ala His Cys Lys
          50          55          60

          Lys Pro Asn Leu Gln Val Phe Leu Gly Lys His Asn Leu Arg Gln Arg
          65          70          75          80

          Glu Ser Ser Gln Glu Gln Ser Ser Val Val Arg Ala Val Ile His Pro
          85          90          95

35      Asp Tyr Asp Ala Ala Ser His Asp Gln Asp Ile Met Leu Leu Arg Leu
          100          105          110

          Ala Arg Pro Ala Lys Leu Ser Glu Leu Ile Gln Pro Leu Pro Leu Glu
          115          120          125

          Arg Asp Cys Ser Ala Asn Thr Thr Ser Cys His Ile Leu Gly Trp Gly
          130          135          140

45      Lys Thr Ala Asp Gly Asp Phe Pro Asp Thr Ile Gln Cys Ala Tyr Ile
          145          150          155          160

          His Leu Val Ser Arg Glu Glu Cys Glu His Ala Tyr Pro Gly Gln Ile
          165          170          175

50      Thr Gln Asn Met Leu Cys Ala Gly Asp Glu Lys Tyr Gly Lys Asp Ser
          180          185          190

          Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gly Asp His Leu Arg
          195          200          205

          Gly Leu Val Ser Trp Gly Asn Ile Pro Cys Gly Ser Lys Glu Lys Pro
          210          215          220

          Gly Val Tyr Thr Asn Val Cys Arg Tyr Thr Asn Trp Ile Gln Lys Thr
          225          230          235          240

65      Ile Gln Ala Lys

```

ES 2 266 245 T3

<210> 26

<211> 277

<212> PRT

5 <213> *Homo sapiens*

<400> 26

10	Met Trp Pro Leu Ala Leu Val Ile Ala Ser Leu Thr Leu Ala Leu Ser	1 5 10 15
	Gly Gly Val Ser Gln Glu Ser Ser Lys Val Leu Asn Thr Asn Gly Thr	20 25 30
15	Ser Gly Phe Leu Pro Gly Gly Tyr Thr Cys Phe Pro His Ser Gln Pro	35 40 45
20	Trp Gln Ala Ala Leu Leu Val Gln Gly Arg Leu Leu Cys Gly Gly Val	50 55 60
	Leu Val His Pro Lys Trp Val Leu Thr Ala Ala His Cys Leu Lys Glu	65 70 75 80
25	Gly Leu Lys Val Tyr Leu Gly Lys His Ala Leu Gly Arg Val Glu Ala	85 90 95
	Gly Glu Gln Val Arg Glu Val Val His Ser Ile Pro His Pro Glu Tyr	100 105 110
30	Arg Arg Ser Pro Thr His Leu Asn His Asp His Asp Ile Met Leu Leu	115 120 125
35	Glu Leu Gln Ser Pro Val Gln Leu Thr Gly Tyr Ile Gln Thr Leu Pro	130 135 140
	Leu Ser His Asn Asn Arg Leu Thr Pro Gly Thr Thr Cys Arg Val Ser	145 150 155 160
40	Gly Trp Gly Thr Thr Thr Ser Pro Gln Val Asn Tyr Pro Lys Thr Leu	165 170 175
	Gln Cys Ala Asn Ile Gln Leu Arg Ser Asp Glu Glu Cys Arg Gln Val	180 185 190
45	Tyr Pro Gly Lys Ile Thr Asp Asn Met Leu Cys Ala Gly Thr Lys Glu	195 200 205
50	Gly Gly Lys Asp Ser Cys Glu Gly Asp Ser Gly Gly Pro Leu Val Cys	210 215 220
	Asn Arg Thr Leu Tyr Gly Ile Val Ser Trp Gly Asp Phe Pro Cys Gly	225 230 235 240
55	Gln Pro Asp Arg Pro Gly Val Tyr Thr Arg Val Ser Arg Tyr Val Leu	245 250 255
60	Trp Ile Arg Glu Thr Ile Arg Lys Tyr Glu Thr Gln Gln Gln Lys Trp	260 265 270
65	Leu Lys Gly Pro Gln	275

ES 2 266 245 T3

<210> 27
 <211> 251
 <212> PRT
 5 <213> *Homo sapiens*
 <400> 27

10 Met Phe Leu Leu Leu Thr Ala Leu Gln Val Leu Ala Ile Ala Met Thr
 1 5 10 15
 Gln Ser Gln Glu Asp Glu Asn Lys Ile Ile Gly Gly His Thr Cys Thr
 20 25 30
 15 Arg Ser Ser Gln Pro Trp Gln Ala Ala Leu Leu Ala Gly Pro Arg Arg
 35 40 45
 Arg Phe Leu Cys Gly Gly Ala Leu Leu Ser Gly Gln Trp Val Ile Thr
 50 55 60
 20 Ala Ala His Cys Gly Arg Pro Ile Leu Gln Val Ala Leu Gly Lys His
 65 70 75 80
 25 Asn Leu Arg Arg Trp Glu Ala Thr Gln Gln Val Leu Arg Val Val Arg
 85 90 95
 Gln Val Thr His Pro Asn Tyr Asn Ser Arg Thr His Asp Asn Asp Leu
 100 105 110
 30 Met Leu Leu Gln Leu Gln Gln Pro Ala Arg Ile Gly Arg Ala Val Arg
 115 120 125
 35 Pro Ile Glu Val Thr Gln Ala Cys Ala Ser Pro Gly Thr Ser Cys Arg
 130 135 140
 Val Ser Gly Trp Gly Thr Ile Ser Ser Pro Ile Ala Arg Tyr Pro Ala
 145 150 155 160
 40 Ser Leu Gln Cys Val Asn Ile Asn Ile Ser Pro Asp Glu Val Cys Gln
 165 170 175
 45 Lys Ala Tyr Pro Arg Thr Ile Thr Pro Gly Met Val Cys Ala Gly Val
 180 185 190
 Pro Gln Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu
 195 200 205
 50 Val Cys Arg Gly Gln Leu Gln Gly Leu Val Ser Trp Gly Met Glu Arg
 210 215 220
 55 Cys Ala Leu Pro Gly Tyr Pro Gly Val Tyr Thr Asn Leu Cys Lys Tyr
 225 230 235 240
 Arg Ser Trp Ile Glu Glu Thr Met Arg Asp Lys
 245 250

60 <210> 28
 <211> 277
 <212> PRT
 65 <213> *Homo sapiens*

ES 2 266 245 T3

<400> 28

5 Met Gln Arg Leu Arg Trp Leu Arg Asp Trp Lys Ser Ser Gly Arg Gly
 1 5 10 15
 Leu Thr Ala Ala Lys Glu Pro Gly Ala Arg Ser Ser Pro Leu Gln Ala
 20 25 30
 10 Met Arg Ile Leu Gln Leu Ile Leu Leu Ala Leu Ala Thr Gly Leu Val
 35 40 45
 Gly Gly Glu Thr Arg Ile Ile Lys Gly Phe Glu Cys Lys Pro His Ser
 50 55 60
 15 Gln Pro Trp Gln Ala Ala Leu Phe Glu Lys Thr Arg Leu Leu Cys Gly
 65 70 75 80
 Ala Thr Leu Ile Ala Pro Arg Trp Leu Leu Thr Ala Ala His Cys Leu
 85 90 95
 20 Lys Pro Arg Tyr Ile Val His Leu Gly Gln His Asn Leu Gln Lys Glu
 100 105 110
 25 Glu Gly Cys Glu Gln Thr Arg Thr Ala Thr Glu Ser Phe Pro His Pro
 115 120 125
 Gly Phe Asn Asn Ser Leu Pro Asn Lys Asp His Arg Asn Asp Ile Met
 130 135 140
 30 Leu Val Lys Met Ala Ser Pro Val Ser Ile Thr Trp Ala Val Arg Pro
 145 150 155 160
 Leu Thr Leu Ser Ser Arg Cys Val Thr Ala Gly Thr Ser Cys Leu Ile
 165 170 175
 40 Ser Gly Trp Gly Ser Thr Ser Ser Pro Gln Leu Arg Leu Pro His Thr
 180 185 190
 Leu Arg Cys Ala Asn Ile Thr Ile Ile Glu His Gln Lys Cys Glu Asn
 195 200 205
 45 Ala Tyr Pro Gly Asn Ile Thr Asp Thr Met Val Cys Ala Ser Val Gln
 210 215 220
 50 Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val
 225 230 235 240
 Cys Asn Gln Ser Leu Gln Gly Ile Ile Ser Trp Gly Gln Asp Pro Cys
 245 250 255
 55 Ala Ile Thr Arg Lys Pro Gly Val Tyr Thr Lys Val Cys Lys Tyr Val
 260 265 270
 60 Asp Trp Ile Gln Glu
 275

<210> 29

<211> 250

65 <212> PRT

<213> *Homo sapiens*

ES 2 266 245 T3

<400> 29

```

Met Lys Leu Gly Leu Leu Cys Ala Leu Leu Ser Leu Leu Ala Gly His
1      5      10      15
5
Gly Trp Ala Asp Thr Arg Ala Ile Gly Ala Glu Glu Cys Arg Pro Asn
20      25      30
10
Ser Gln Pro Trp Gln Ala Gly Leu Phe His Leu Thr Arg Leu Phe Cys
35      40      45
15
Gly Ala Thr Leu Ile Ser Asp Arg Trp Leu Leu Thr Ala Ala His Cys
50      55      60
20
Arg Lys Pro Tyr Leu Trp Val Arg Leu Gly Glu His His Leu Trp Lys
65      70      75      80
25
Trp Glu Gly Pro Glu Gln Leu Phe Arg Val Thr Asp Phe Phe Pro His
85      90      95
30
Pro Gly Phe Asn Lys Asp Leu Ser Ala Asn Asp His Asn Asp Asp Ile
100     105     110
35
Met Leu Ile Arg Leu Pro Arg Gln Ala Arg Leu Ser Pro Ala Val Gln
115     120     125
40
Pro Leu Asn Leu Ser Gln Thr Cys Val Ser Pro Gly Met Gln Cys Leu
130     135     140
45
Ile Ser Gly Trp Gly Ala Val Ser Ser Pro Lys Ala Leu Phe Pro Val
145     150     155     160
50
Thr Leu Gln Cys Ala Asn Ile Ser Ile Leu Glu Asn Lys Leu Cys His
165     170     175
55
Trp Ala Tyr Pro Gly His Ile Ser Asp Ser Met Leu Cys Ala Gly Leu
180     185     190
60
Trp Glu Gly Gly Arg Gly Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu
195     200     205
65
Val Cys Asn Gly Thr Leu Ala Gly Val Val Ser Gly Gly Ala Glu Pro
210     215     220
70
Cys Ser Arg Pro Arg Arg Pro Ala Val Tyr Thr Ser Val Cys His Tyr
225     230     235     240
75
Leu Asp Trp Ile Gln Glu Ile Met Glu Asn
245     250

```

<210> 30

<211> 276

<212> PRT

<213> *Homo sapiens*

ES 2 266 245 T3

<400> 30

```

Met Arg Ala Pro His Leu His Leu Ser Ala Ala Ser Gly Ala Arg Ala
1      5      10      15

5  Leu Ala Lys Leu Leu Pro Leu Leu Met Ala Gln Leu Trp Ala Ala Glu
    20      25      30

10 Ala Ala Leu Leu Pro Gln Asn Asp Thr Arg Leu Asp Pro Glu Ala Tyr
    35      40      45

Gly Ala Pro Cys Ala Arg Gly Ser Gln Pro Trp Gln Val Ser Leu Phe
50      55      60

15 Asn Gly Leu Ser Phe His Cys Ala Gly Val Leu Val Asp Gln Ser Trp
65      70      75      80

Val Leu Thr Ala Ala His Cys Gly Asn Lys Pro Leu Trp Ala Arg Val
85      90      95

20 Gly Asp Asp His Leu Leu Leu Leu Gln Gly Glu Gln Leu Arg Arg Thr
    100      105      110

25 Thr Arg Ser Val Val His Pro Lys Tyr His Gln Gly Ser Gly Pro Ile
    115      120      125

Leu Pro Arg Arg Thr Asp Glu His Asp Leu Met Leu Leu Lys Leu Ala
130      135      140

30 Arg Pro Val Val Pro Gly Pro Arg Val Arg Ala Leu Gln Leu Pro Tyr
145      150      155      160

35 Arg Cys Ala Gln Pro Gly Asp Gln Cys Gln Val Ala Gly Trp Gly Thr
    165      170      175

Thr Ala Ala Arg Arg Val Lys Tyr Asn Lys Gly Leu Thr Cys Ser Ser
180      185      190

40 Ile Thr Ile Leu Ser Pro Lys Glu Cys Glu Val Phe Tyr Pro Gly Val
195      200      205

45 Val Thr Asn Asn Met Ile Cys Ala Gly Leu Asp Arg Gly Gln Asp Pro
210      215      220

Cys Gln Ser Asp Ser Gly Gly Pro Leu Val Cys Asp Glu Thr Leu Gln
225      230      235      240

50 Gly Ile Leu Ser Trp Gly Val Tyr Pro Cys Gly Ser Ala Gln His Pro
    245      250      255

Ala Val Tyr Thr Gln Ile Cys Lys Tyr Met Ser Trp Ile Asn Lys Val
260      265      270

Ile Arg Ser Asn
275

```

<210> 31

<211> 248

<212> PRT

<213> *Homo sapiens*

ES 2 266 245 T3

<400> 31

Met Gly Leu Ser Ile Phe Leu Leu Leu Cys Val Leu Gly Leu Ser Gln
1 5 10 15

Ala Ala Thr Pro Lys Ile Phe Asn Gly Thr Glu Cys Gly Arg Asn Ser
20 25 30

Gln Pro Trp Gln Val Gly Leu Phe Glu Gly Thr Ser Leu Arg Cys Gly
35 40 45

Gly Val Leu Ile Asp His Arg Trp Val Leu Thr Ala Ala His Cys Ser
50 55 60

Gly Ser Arg Tyr Trp Val Arg Leu Gly Glu His Ser Leu Ser Gln Leu
65 70 75 80

Asp Trp Thr Glu Gln Ile Arg His Ser Gly Phe Ser Val Thr His Pro
85 90 95

Gly Tyr Leu Gly Ala Ser Thr Ser His Glu His Asp Leu Arg Leu Leu
100 105 110

Arg Leu Arg Leu Pro Val Arg Val Thr Ser Ser Val Gln Pro Leu Pro
115 120 125

Leu Pro Asn Asp Cys Ala Thr Ala Gly Thr Glu Cys His Val Ser Gly
130 135 140

Trp Gly Ile Thr Asn His Pro Arg Asn Pro Phe Pro Asp Leu Leu Gln
145 150 155 160

Cys Leu Asn Leu Ser Ile Val Ser His Ala Thr Cys His Gly Val Tyr
165 170 175

Pro Gly Arg Ile Thr Ser Asn Met Val Cys Ala Gly Gly Val Pro Gly
180 185 190

Gln Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gly Gly
195 200 205

Val Leu Gln Gly Leu Val Ser Trp Gly Ser Val Gly Pro Cys Gly Gln
210 215 220

Asp Gly Ile Pro Gly Val Tyr Thr Tyr Ile Cys Lys Tyr Val Asp Trp
225 230 235 240

Ile Arg Met Ile Met Arg Asn Asn
245

<210> 32

<211> 260

<212> PRT

<213> *Homo sapiens*

ES 2 266 245 T3

<400> 32

Met Gly Arg Pro Arg Pro Arg Ala Ala Lys Thr Trp Met Phe Leu Leu
1 5 10 15

Leu Leu Gly Gly Ala Trp Ala Gly His Ser Arg Ala Gln Glu Asp Lys
20 25 30

Val Leu Gly Gly His Glu Cys Gln Pro His Ser Gln Pro Trp Gln Ala
35 40 45

Ala Leu Phe Gln Gly Gln Gln Leu Leu Cys Gly Gly Val Leu Val Gly
50 55 60

Gly Asn Trp Val Leu Thr Ala Ala His Cys Lys Lys Pro Lys Tyr Thr
65 70 75 80

Val Arg Leu Gly Asp His Ser Leu Gln Asn Lys Asp Gly Pro Glu Gln
85 90 95

Glu Ile Pro Val Val Gln Ser Ile Pro His Pro Cys Tyr Asn Ser Ser
100 105 110

Asp Val Glu Asp His Asn His Asp Leu Met Leu Leu Gln Leu Arg Asp
115 120 125

Gln Ala Ser Leu Gly Ser Lys Val Lys Pro Ile Ser Leu Ala Asp His
130 135 140

Cys Thr Gln Pro Gly Gln Lys Cys Thr Val Ser Gly Trp Gly Thr Val
145 150 155 160

Thr Ser Pro Arg Glu Asn Phe Pro Asp Thr Leu Asn Cys Ala Glu Val
165 170 175

Lys Ile Phe Pro Gln Lys Lys Cys Glu Asp Ala Tyr Pro Gly Gln Ile
180 185 190

Thr Asp Gly Met Val Cys Ala Gly Ser Ser Lys Gly Ala Asp Thr Cys
195 200 205

Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Asp Gly Ala Leu Gln Gly
210 215 220

Ile Thr Ser Trp Gly Ser Asp Pro Cys Gly Arg Ser Asp Lys Pro Gly
225 230 235 240

Val Tyr Thr Asn Ile Cys Arg Tyr Leu Asp Trp Ile Lys Lys Ile Ile
245 250 255

Gly Ser Lys Gly
260

<210> 33

<211> 261

<212> PRT

<213> *Homo sapiens*

ES 2 266 245 T3

<400> 33

Met Trp Val Pro Val Val Phe Leu Thr Leu Ser Val Thr Trp Ile Gly
1 5 10 15

Ala Ala Pro Leu Ile Leu Ser Arg Ile Val Gly Gly Trp Glu Cys Glu
20 25 30

Lys His Ser Gln Pro Trp Gln Val Leu Val Ala Ser Arg Gly Arg Ala
35 40 45

Val Cys Gly Gly Val Leu Val His Pro Gln Trp Val Leu Thr Ala Ala
50 55 60

His Cys Ile Arg Asn Lys Ser Val Ile Leu Leu Gly Arg His Ser Leu
65 70 75 80

Phe His Pro Glu Asp Thr Gly Gln Val Phe Gln Val Ser His Ser Phe
85 90 95

Pro His Pro Leu Tyr Asp Met Ser Leu Leu Lys Asn Arg Phe Leu Arg
100 105 110

Pro Gly Asp Asp Ser Ser His Asp Leu Met Leu Leu Arg Leu Ser Glu
115 120 125

Pro Ala Glu Leu Thr Asp Ala Val Lys Val Met Asp Leu Pro Thr Gln
130 135 140

Glu Pro Ala Leu Gly Thr Thr Cys Tyr Ala Ser Gly Trp Gly Ser Ile
145 150 155 160

Glu Pro Glu Glu Phe Leu Thr Pro Lys Lys Leu Gln Cys Val Asp Leu
165 170 175

His Val Ile Ser Asn Asp Val Cys Ala Gln Val His Pro Gln Lys Val
180 185 190

Thr Lys Phe Met Leu Cys Ala Gly Arg Trp Thr Gly Gly Lys Ser Thr
195 200 205

Cys Ser Gly Asp Ser Gly Gly Pro Leu Val Cys Asn Gly Val Leu Gln
210 215 220

Gly Ile Thr Ser Trp Gly Ser Glu Pro Cys Ala Leu Pro Glu Arg Pro
225 230 235 240

Ser Leu Tyr Thr Lys Val Val His Tyr Arg Lys Trp Ile Lys Asp Thr
245 250 255

Ile Val Ala Asn Pro
260

<210> 34

<211> 261

<212> PRT

<213> *Homo sapiens*

ES 2 266 245 T3

<400> 34

```

Met Trp Asp Leu Val Leu Ser Ile Ala Leu Ser Val Gly Cys Thr Gly
1      5      10      15

5
Ala Val Pro Leu Ile Gln Ser Arg Ile Val Gly Gly Trp Glu Cys Glu
20      25      30

10
Lys His Ser Gln Pro Trp Gln Val Ala Val Tyr Ser His Gly Trp Ala
35      40      45

15
His Cys Gly Gly Val Leu Val His Pro Gln Trp Val Leu Thr Ala Ala
50      55      60

His Cys Leu Lys Lys Asn Ser Gln Val Trp Leu Gly Arg His Asn Leu
65      70      75      80

20
Phe Glu Pro Glu Asp Thr Gly Gln Arg Val Pro Val Ser His Ser Phe
85      90      95

Pro His Pro Leu Tyr Asn Met Ser Leu Leu Lys His Gln Ser Leu Arg
100     105     110

25
Pro Asp Glu Asp Ser Ser His Asp Leu Met Leu Leu Arg Leu Ser Glu
115     120     125

30
Pro Ala Lys Ile Thr Asp Val Val Lys Val Leu Gly Leu Pro Thr Gln
130     135     140

Glu Pro Ala Leu Gly Thr Thr Cys Tyr Ala Ser Gly Trp Gly Ser Ile
145     150     155     160

35
Glu Pro Glu Glu Phe Leu Arg Pro Arg Ser Leu Gln Cys Val Ser Leu
165     170     175

His Leu Leu Ser Asn Asp Met Cys Ala Arg Ala Tyr Ser Glu Lys Val
180     185     190

40
Thr Glu Phe Met Leu Cys Ala Gly Leu Trp Thr Gly Gly Lys Asp Thr
195     200     205

45
Cys Gly Gly Asp Ser Gly Gly Pro Leu Val Cys Asn Gly Val Leu Gln
210     215     220

Gly Ile Thr Ser Trp Gly Pro Glu Pro Cys Ala Leu Pro Glu Lys Pro
225     230     235     240

50
Ala Val Tyr Thr Lys Val Val His Tyr Arg Lys Trp Ile Lys Asp Thr
245     250     255

55
Ile Ala Ala Asn Pro
260

```

<210> 35

<211> 262

60 <212> PRT

<213> *Homo sapiens*

<400> 35

```

Met Trp Phe Leu Val Leu Cys Leu Ala Leu Ser Leu Gly Gly Thr Gly
1      5      10      15

```

ES 2 266 245 T3

Ala Ala Pro Pro Ile Gln Ser Arg Ile Val Gly Gly Trp Glu Cys Glu
 20 25 30
 5 Gln His Ser Gln Pro Trp Gln Ala Ala Leu Tyr His Phe Ser Thr Phe
 35 40 45
 Gln Cys Gly Gly Ile Leu Val His Arg Gln Trp Val Leu Thr Ala Ala
 50 55 60
 10 His Cys Ile Ser Asp Asn Tyr Gln Leu Trp Leu Gly Arg His Asn Leu
 65 70 75 80
 Phe Asp Asp Glu Asn Thr Ala Gln Phe Val His Val Ser Glu Ser Phe
 85 90 95
 15 Pro His Pro Gly Phe Asn Met Ser Leu Leu Glu Asn His Thr Arg Gln
 100 105 110
 20 Ala Asp Glu Asp Tyr Ser His Asp Leu Met Leu Leu Arg Leu Thr Glu
 115 120 125
 Pro Ala Asp Thr Ile Thr Asp Ala Val Lys Val Val Glu Leu Pro Thr
 130 135 140
 25 Glu Glu Pro Glu Val Gly Ser Thr Cys Leu Ala Ser Gly Trp Gly Ser
 145 150 155 160
 30 Ile Glu Pro Glu Asn Phe Ser Phe Pro Asp Asp Leu Gln Cys Val Asp
 165 170 175
 35 Leu Lys Ile Leu Pro Asn Asp Glu Cys Lys Lys Ala His Val Gln Lys
 180 185 190
 Val Thr Asp Phe Met Leu Cys Val Gly His Leu Glu Gly Gly Lys Asp
 195 200 205
 40 Thr Cys Val Gly Asp Ser Gly Gly Pro Leu Met Cys Asp Gly Val Leu
 210 215 220
 Gln Gly Val Thr Ser Trp Gly Tyr Val Pro Cys Gly Thr Pro Asn Lys
 225 230 235 240
 45 Pro Ser Val Ala Val Arg Val Leu Ser Tyr Val Lys Trp Ile Glu Asp
 245 250 255
 50 Thr Ile Ala Glu Asn Ser
 260

<210> 36

<211> 293

<212> PRT

<213> *Homo sapiens*

ES 2 266 245 T3

<400> 36

5

Met Ala Thr Ala Arg Pro Pro Trp Met Trp Val Leu Cys Ala Leu Ile
1 5 10 15

Thr Ala Leu Leu Leu Gly Val Thr Glu His Val Leu Ala Asn Asn Asp
20 25 30

10

Val Ser Cys Asp His Pro Ser Asn Thr Val Pro Ser Gly Ser Asn Gln
35 40 45

Asp Leu Gly Ala Gly Ala Gly Glu Asp Ala Arg Ser Asp Asp Ser Ser
50 55 60

15

Ser Arg Ile Ile Asn Gly Ser Asp Cys Asp Met His Thr Gln Pro Trp
65 70 75 80

20

Gln Ala Ala Leu Leu Leu Arg Pro Asn Gln Leu Tyr Cys Gly Ala Val
85 90 95

Leu Val His Pro Gln Trp Leu Leu Thr Ala Ala His Cys Arg Lys Lys
100 105 110

25

Val Phe Arg Val Arg Leu Gly His Tyr Ser Leu Ser Pro Val Tyr Glu
115 120 125

30

Ser Gly Gln Gln Met Phe Gln Gly Val Lys Ser Ile Pro His Pro Gly
130 135 140

Tyr Ser His Pro Gly His Ser Asn Asp Leu Met Leu Ile Lys Leu Asn
145 150 155 160

35

Arg Arg Ile Arg Pro Thr Lys Asp Val Arg Pro Ile Asn Val Ser Ser
165 170 175

40

His Cys Pro Ser Ala Gly Thr Lys Cys Leu Val Ser Gly Trp Gly Thr
180 185 190

Thr Lys Ser Pro Gln Val His Phe Pro Lys Val Leu Gln Cys Leu Asn
195 200 205

45

Ile Ser Val Leu Ser Gln Lys Arg Cys Glu Asp Ala Tyr Pro Arg Gln
210 215 220

File Asp Asp Thr Met Phe Cys Ala Gly Asp Lys Ala Gly Arg Asp Ser
225 230 235 240

50

Cys Gln Gly Asp Ser Gly Gly Pro Val Val Cys Asn Gly Ser Leu Gln
245 250 255

55

Gly Leu Val Ser Trp Gly Asp Tyr Pro Cys Ala Arg Pro Asn Arg Pro
260 265 270

Gly Val Tyr Thr Asn Leu Cys Lys Phe Thr Lys Trp Ile Gln Glu Thr
275 280 285

60

Ile Gln Ala Asn Ser
290

<210> 37

65 <211> 253

<212> PRT

<213> *Homo sapiens*

ES 2 266 245 T3

<400> 37

Met Ala Thr Ala Gly Asn Pro Trp Gly Trp Phe Leu Gly Tyr Leu Ile
1 5 10 15

Leu Gly Val Ala Gly Ser Leu Val Ser Gly Ser Cys Ser Gln Ile Ile
20 25 30

Asn Gly Glu Asp Cys Ser Pro His Ser Gln Pro Trp Gln Ala Ala Leu
35 40 45

Val Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln
50 55 60

Trp Val Leu Ser Ala Ala His Cys Phe Gln Asn Ser Tyr Thr Ile Gly
65 70 75 80

Leu Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met
85 90 95

Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro Leu
100 105 110

Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Ser
115 120 125

Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly
130 135 140

Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly Arg Met
145 150 155 160

Pro Thr Val Leu Gln Cys Val Asn Val Ser Val Val Ser Glu Glu Val
165 170 175

Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe Cys Ala
180 185 190

Gly Gly Gly His Asp Gln Lys Asp Ser Cys Asn Gly Asp Ser Gly Gly
195 200 205

Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe Gly Lys
210 215 220

Ala Pro Cys Gly Gln Val Gly Val Pro Gly Val Tyr Thr Asn Leu Cys
225 230 235 240

Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
245 250

<210> 38

<211> 257

<212> PRT

<213> *Homo sapiens*

ES 2 266 245 T3

<400> 38

```

Met Ala Arg Ser Leu Leu Pro Leu Gln Ile Leu Leu Leu Ser Leu
1      5      10      15

5  Ala Leu Glu Thr Ala Gly Glu Glu Ala Gln Gly Asp Lys Ile Ile Asp
    20      25      30

10 Gly Ala Pro Cys Ala Arg Gly Ser His Pro Trp Gln Val Ala Leu Leu
    35      40      45

15 Ser Gly Asn Gln Leu His Cys His Ser Cys Cys Glu Gly Gly Val Leu
    50      55      60

15 Val Asn Glu Arg Trp Val Leu Thr Ala Ala His Cys Lys Met Asn Glu
    65      70      75      80

20 Tyr Thr Val His Leu Gly Ser Asp Thr Leu Gly Asp Arg Arg Ala Gln
    85      90      95

25 Arg Ile Lys Ala Ser Lys Ser Phe Arg His Pro Gly Tyr Ser Thr Gln
    100     105     110

30 Thr His Val Asn Asp Leu Met Leu Val Lys Leu Asn Ser Gln Ala Arg
    115     120     125

30 Leu Ser Ser Met Val Lys Lys Val Arg Leu Pro Ser Arg Cys Glu Pro
    130     135     140

35 Pro Gly Thr Thr Cys Thr Val Ser Gly Trp Gly Thr Thr Thr Ser Pro
    145     150     155     160

35 Asp Val Thr Phe Pro Asp Leu Met Cys Val Asp Val Lys Leu Ile Ser
    165     170     175

40 Pro Gln Asp Cys Thr Lys Val Tyr Lys Asp Leu Leu Glu Asn Ser Met
    180     185     190

40 Leu Cys Ala Gly Ile Pro Asp Ser Lys Lys Asn Ala Cys Asn Gly Asp
    195     200     205

45 Ser Gly Gly Pro Leu Val Cys Arg Gly Thr Leu Gln Gly Leu Val Ser
    210     215     220

50 Trp Gly Thr Phe Pro Cys Gly Gln Pro Asn Asp Pro Gly Val Tyr Thr
    225     230     235     240

50 Gln Val Cys Lys Phe Thr Lys Trp Ile Asn Asp Thr Met Lys Lys His
    245     250     255

```

Arg

<210> 39

<211> 24

<212> ADN

<213> Secuencia artificial

<220>

<223> segmentos/péptidos procedentes de secuencia humana

<400> 39

cacaacgagc ctgggaccgc tggg

24

ES 2 266 245 T3

	<210> 40	
	<211> 6	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> segmentos/péptidos procedentes de secuencia humana	
10	<400> 40	
	attaaa	6
15	<210> 41	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> segmentos/péptidos procedentes de secuencia humana	
	<400> 41	
25	atccctccat tcccatcttt	20
	<210> 42	
	<211> 20	
30	<212> ADN	
	<213> Secuencia artificial	
	<220>	
35	<223> segmentos/péptidos procedentes de secuencia humana	
	<400> 42	
	cacatacaat tctctggttc	20
40	<210> 43	
	<211> 20	
	<212> ADN	
45	<213> Secuencia artificial	
	<220>	
	<223> segmentos/péptidos procedentes de secuencia humana	
	<400> 43	
50	agtgacactg ttcagaatt	20
	<210> 44	
55	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
60	<223> segmentos/péptidos procedentes de secuencia humana	
	<400> 44	
65	ccccaatctc acgagtgcac	20
	<210> 45	
	<211> 20	

ES 2 266 245 T3

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> segmentos/péptidos procedentes de secuencia humana	
	<400> 45	
	gtcggctctg gagacatttc	20
10	<210> 46	
	<211> 19	
	<212> ADN	
15	<213> Secuencia artificial	
	<220>	
	<223> segmentos/péptidos procedentes de secuencia humana	
20	<400> 46	
	aactggggag gcttgagtc	19
25	<210> 47	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> segmentos/péptidos procedentes de secuencia humana	
	<400> 47	
35	ctccttctg ctggcatcca	20
	<210> 48	
	<211> 20	
40	<212> ADN	
	<213> Secuencia artificial	
	<220>	
45	<223> segmentos/péptidos procedentes de secuencia humana	
	<400> 48	
	atcacacggg tggcatgtg	20
50	<210> 49	
	<211> 20	
	<212> ADN	
55	<213> Secuencia artificial	
	<220>	
	<223> segmentos/péptidos procedentes de secuencia humana	
60	<400> 49	
	caagtggctc tctacgagcg	20
65	<210> 50	
	<211> 20	
	<212> ADN	

ES 2 266 245 T3

<213> Secuencia artificial

<220>

<223> segmentos/péptidos procedentes de secuencia humana

5 <400> 50

gacaccaggc ttggtggtgt

20

10

15

20

25

30

35

40

45

50

55

60

65