



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 305 116**

51 Int. Cl.:
C07K 14/47 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **01980249 .5**

86 Fecha de presentación : **16.08.2001**

87 Número de publicación de la solicitud: **1313766**

87 Fecha de publicación de la solicitud: **28.05.2003**

54 Título: **Coactivación de receptores nucleares.**

30 Prioridad: **21.08.2000 EP 00202905**
14.05.2001 EP 01201771

45 Fecha de publicación de la mención BOPI:
01.11.2008

45 Fecha de la publicación del folleto de la patente:
01.11.2008

73 Titular/es: **N.V. Organon**
Kloosterstraat 6
5349 AB Oss, NL

72 Inventor/es: **Dechering, Koen, Jacob y**
Mosselman, Sietse

74 Agente: **Ungría López, Javier**

ES 2 305 116 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Coactivación de receptores nucleares.

5 Esta invención se refiere a una nueva proteína, a un ensayo para la unión entre un receptor nuclear sensible y un coactivador y a un método de modulación de la actividad transcripcional promovida por un receptor nuclear sensible.

Esta invención está en el campo del uso de funciones de receptores nucleares en organismos. Los receptores nucleares tienen un gran intervalo de funciones, pero su papel más destacado consiste en permitir a las células responder a hormonas, tales como estradiol, cortisol y progesterona. Mediante los receptores nucleares, las hormonas influyen en la transcripción de genes en las células, mediante la cual las células pueden modificar su función. Además, se sabe que la activación de la transcripción génica por receptores nucleares está facilitada o incluso permitida por coactivadores. Se conocen varios de dichos coactivadores. Por ejemplo, Nishikawa *et al* (Toxicol. Appl. Pharmacol. 154, págs. 76-83, 1999) describen métodos de exploración para químicos con actividades hormonales usando la interacción de receptores hormonales nucleares con un coactivador. Su método se basa en el descubrimiento de que la interacción entre el coactivador y un receptor nuclear, que es sensible a ese coactivador en particular, puede estabilizarse por un ligando que se une al receptor nuclear. Dichos ensayos también son importantes para desenmarañar la acción de hormonas y los mecanismos para el control de la transcripción de genes en general. Con estas técnicas pueden desarrollarse nuevas medicinas para influir específicamente en los procesos fisiológicos relacionados con el funcionamiento de receptores nucleares para fines terapéuticos, de diagnóstico, cosméticos y anticonceptivos. Esta invención pone a disposición una proteína que tiene la secuencia de aminoácidos que se proporciona en las SEC ID 7, 8, 11 ó 12. Dicha proteína tiene el efecto de coactivar un receptor nuclear sensible en su función promotora de la transcripción de ADN. Con el descubrimiento de la función de dicha proteína como coactivador, esta invención también proporciona un nuevo método de modulación de la actividad transcripcional promovida por un receptor nuclear sensible y un coactivador en un sistema, que comprende la adición de un agente al sistema, interfiriendo el agente con, o potenciando, la función del coactivador de esta invención. Este método puede realizarse mediante la introducción del gen para el coactivador en el sistema para producir su expresión, así como mediante la identificación y posterior uso de compuestos que pueden modificar el funcionamiento del coactivador y/o del complejo coactivador-receptor nuclear. Por lo tanto, el término agente se refiere en la presente memoria, por ejemplo, a un compuesto químico o a un vector, comprendiendo el vector el gen que codifica la proteína de la invención. Dicha modulación puede usarse para interferir con, o potenciar, los procesos de transcripción nuclear con fines terapéuticos, anticonceptivos, de diagnóstico y cosméticos, y puede usarse en ensayos para la cuantificación de la modulación. Por lo tanto, el término sistema se refiere a un entorno definido molecularmente y/o biológicamente y se refiere a sistemas *in vitro* así como a sistemas *in vivo* en los que puede introducirse el agente modulador. Los sistemas *in vitro* son mezclas que permiten la transcripción o traducción de ADN y/o ARN y los sistemas *in vivo* son células, que pueden estar en un cultivo de células o de tejidos o en el interior de un cuerpo humano o animal. Obviamente, la expresión modulación de la actividad de un receptor nuclear puede referirse a un aumento o a una disminución en la actividad transcripcional del receptor, mientras que el término coactivación se refiere a la potenciación de la actividad transcripcional del receptor.

Una secuencia, que supuestamente representa una proteína de este tipo, se ha descrito previamente con el número de gen KIAA0576 en una publicación por Nagase *et al*. (DNA Research vol. 5, págs. 31-39, 1998) con referencia al número de acceso AB011148 para las bases de datos EMBL/GenBank/DDBJ, donde está ligada a la id de proteína BAA25502. El gen KIAA0576 no se ha expresado más allá de la fase de ARN y no se le ha supuesto ninguna otra función más específica que la dirección de ácidos nucleicos en base a una homología muy débil (<30% de nucleótidos idénticos) con una secuencia de ADN conocida.

Puesto que es obvio que modificaciones minoritarias en la secuencia de la proteína son igualmente útiles para el uso descrito anteriormente, la invención también describe una proteína que tiene una secuencia de aminoácidos que tiene al menos una similitud del 90%, o preferiblemente de al menos el 95%, más preferiblemente de al menos el 99% y, más preferiblemente, del 100% con la secuencia de las SEC ID 7, 8, 11 ó 12. El término similitud se refiere a un grado de similitud entre proteína en vista de diferencias en aminoácidos, pero cuyos aminoácidos diferentes son funcionalmente similares en vista de un tamaño, lipofilicidad, acidez, etc. casi iguales. Un porcentaje de similitud se calcula mediante el alineamiento óptimo de las secuencias usando una matriz de puntuación de similitud tal como la matriz blosum62 descrita en Henikoff y Henikoff (Proc. Natl. Acad. Sci. USA 89; 10915-10919 (1991)). El cálculo del porcentaje de similitud y el alineamiento óptimo de dos secuencias, usando la matriz de similitud blosum62 y el algoritmo de Needleman y Wunsch (J. Mol. Biol. 48; 443-453 (1970)), puede realizarse usando el programa GAP del Genetics Computer Group (GCG, Madison, Wisconsin, Estados Unidos) usando los parámetros por defecto del programa.

Un aspecto adicional de la invención es que la función de coactivación descrita anteriormente también puede obtenerse con una proteína que tiene la secuencia de los aminoácidos número 1 a 234 en la SEC ID 7 u 8. Se sabe que las proteínas truncadas pueden tener la misma o parcialmente la misma función que la proteína de mayor tamaño. Se descubrió que también puede usarse esta proteína truncada y, por lo tanto, este aspecto de la invención también proporciona una proteína que tiene una secuencia que tiene una similitud de al menos el 90%, preferiblemente de al menos el 95%, más preferiblemente de al menos el 99% y, más preferiblemente, del 100% con la secuencia de los aminoácidos número 1 a 234 en la SEC ID 7 u 8.

Además, la invención describe una proteína que es una variante de origen natural de una proteína que tiene la secuencia como en las SEC ID 7, 8, 11 ó 12. Proteínas, que tienen una similitud de al menos el 80%, preferiblemente de al menos el 90%, más preferiblemente de al menos el 95%, aún más preferiblemente de al menos el 99% y, más preferiblemente, una similitud del 100% con la proteína definida en las SEC ID 7, 8, 11 ó 12 y que procede de tejidos humanos o de tejidos de especies no humanas, pueden usarse igualmente para la coactivación de un receptor nuclear sensible. Una proteína de origen natural, que pertenece a la clase definida por cosas en común en la estructura y función indicadas de la proteína que se define por las SEC ID 7, 8, 11 ó 12, se denominará mediante la secuencia de letras arbitraria de proteínas COASTER. En documentos anteriores, al menos en documentos que han circulado por nuestro organismo de investigación, la secuencia de letras arbitraria MFC se ha usado para proporcionar un nombre para las proteínas de acuerdo con esta invención. Por lo tanto, MFC y COASTER son sinónimos. Otros fragmentos COASTER también aparecen en tejidos biológicos. Dichos fragmentos están codificados por variantes de corte y empalme.

Además, habiendo descrito la estructura de la proteína y la función de una proteína COASTER, la invención también describe una proteína que es un equivalente funcional homólogo de la proteína como se ha definido anteriormente. Un equivalente funcional homólogo es una proteína que no cumple completamente las características definidas anteriormente porque tiene una similitud inferior a uno de los porcentajes indicados anteriormente, o porque no es el fragmento que se ha definido anteriormente, o porque no es de origen natural, pero que todavía tiene una función característica en común con las proteínas de la invención definidas anteriormente. Por ejemplo, un fragmento de una proteína COASTER no sólo puede usarse en solitario para la coactivación del receptor nuclear sensible, sino que también puede fusionarse con otras secuencias proteicas conocidas de otras proteínas para beneficiarse del efecto activador del receptor nuclear. Dichos fragmentos no necesitan tener una identidad del 100% con las secuencias de los fragmentos de proteínas COASTER intactas. Además, pueden generarse artificialmente proteínas funcionalmente equivalentes a proteínas COASTER y variantes de corte y empalme de las mismas mediante mutaciones, inserciones o delecciones deliberadas en la secuencia polinucleotídica codificante. Un alto grado de similitud de una proteína con una proteína específicamente caracterizada en esta descripción hace a una proteína funcionalmente equivalente a una proteína COASTER.

En vista de la utilidad de las proteínas, un aspecto de la invención es proporcionar el uso de un polinucleótido que codifica la proteína de esta invención produciendo la expresión del polinucleótido en un sistema que también comprenda un receptor nuclear sensible. El polinucleótido puede ser ADN o ARN puesto que tanto una secuencia de nucleótidos de ADN como una secuencia de ARN servirá con el fin de codificar las proteínas de la invención. Son ADN preferidos para usar en esta invención las secuencias que se proporcionan en las SEC ID 1, 2, 3, 4, 5, 6, 9 ó 10. Puede obtenerse la expresión incorporando artificialmente el polinucleótido en células eucariotas, bacterias, plásmidos o vectores y permitiendo la transcripción del gen por métodos generalmente conocidos en la técnica. El sistema, que puede ser una mezcla con componentes para la transcripción y/o traducción de ADN o ARN o una célula intacta, posiblemente parte de un organismo completo, también debería contener o generar el receptor nuclear sensible para ser capaz de usar la interacción entre la proteína de la invención y el receptor nuclear. Una fuente práctica de información sobre estas técnicas puede encontrarse en Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989.

En el contexto de nucleótidos también se puede hacer referencia a grados de similitud, en los que tripletes de nucleótidos que codifican el mismo aminoácido pueden reemplazar a tripletes en las secuencias anteriores. Es más habitual referirse a secuencias de nucleótidos homólogas:

La homología de un polinucleótido se define como:

$$\text{Homología (\%)} = \frac{\text{Número de restos idénticos entre dos secuencias}}{\text{Longitud de secuencias alineadas} - \text{Longitud de todos los huecos}}$$

después del alineamiento óptimo de las secuencias. Puede realizarse un alineamiento óptimo usando el algoritmo de Needleman y Wunsch (J. Mol. Biol. 48; 443-453 (1970)) que maximiza el número de emparejamientos y minimiza el número de huecos. De hecho, puede definirse un intervalo más reducido de proteínas mediante el reemplazo del término similitud por homología, que tiene la definición análoga a la de homología de un polinucleótido. Dicho uso del término homología para una proteína de la invención se basa en el porcentaje de identidad de restos aminoácidos en la proteína. Como alternativa, la expresión grado de homología de una proteína, definida de esta forma, también puede reemplazarse por la expresión grado de identidad de una proteína.

Para adaptar la variabilidad de codones, la invención también incluye secuencias polinucleotídicas que codifican para las mismas secuencias de aminoácidos que las secuencias descritas en la presente memoria. La información de secuencias, como se proporciona en la presente memoria, no debería interpretarse tan limitadamente como para que sea necesaria la exclusión de bases identificadas erróneamente. La secuencia específica descrita en la presente memoria puede usarse fácilmente para aislar los genes completos de varias otras especies o variantes alélicas. Por ejemplo, la secuencia puede usarse para preparar sondas o como una fuente para preparar oligonucleótidos sintéticos para usarlos como cebadores en reacciones de amplificación de ADN que permiten el aislamiento y la identificación de los genes variantes completos. La secuencia genética completa puede usarse en la preparación de moléculas de vector para la expresión de la proteína en células hospedadoras adecuadas.

La presente invención se refiere además a polinucleótidos que tienen ligeras variaciones o que tienen sitios polimórficos. Los polinucleótidos que tienen ligeras variaciones pueden codificar polipéptidos variantes que conserven la misma función o actividad biológica que la proteína madura natural. Los sitios polimórficos son útiles con fines de diagnóstico. En otro aspecto, la invención proporciona un método para aislar un polinucleótido que comprende las etapas de: a) hibridar un ADN de acuerdo con la presente invención en condiciones rigurosas frente a ácidos nucleicos que sean ARN, ADN (genómico) o ADNc aislados preferiblemente a partir de tejidos que expresen un nivel elevado del polinucleótido de interés; y b) aislar dichos ácidos nucleicos por métodos conocidos por un especialista en la técnica. Los tejidos son preferiblemente de origen humano. Preferiblemente, se aíslan ácidos ribonucleicos de oocitos, ovarios o testículos. Las condiciones de hibridación son preferiblemente muy rigurosas. De acuerdo con la presente invención, el término "rigurosa" se refiere a condiciones de lavado de SSC 1x, SDS al 0,1% a una temperatura de 65°C; las condiciones muy rigurosas se refieren a una reducción de SSC hasta SSC 0,3x, más preferiblemente hasta SSC 0,1x. Preferiblemente, los dos primeros lavados se llevan a cabo posteriormente dos veces, cada una durante 15-30 minutos. Si existe la necesidad de lavar en condiciones muy rigurosas se realiza un lavado adicional con SSC 0,1x una vez durante 15 minutos. La hibridación puede realizarse, por ejemplo, durante una noche en tampón fosfato 0,5 M a pH 7,5/SDS al 7% a 65°C. Como una alternativa el método para aislar el gen puede comprender metodología de amplificación de genes usando cebadores obtenidos a partir del ácido nucleico de acuerdo con la invención. También pueden obtenerse ADNc completos mediante la combinación de clones obtenidos, por ejemplo, por hibridación con clones de ADNc de RACE, por ejemplo.

Para usar la modulación de la transcripción el sistema también debería comprender genes para receptores nucleares sensibles y elementos promotores sensibles para estos receptores nucleares. El polinucleótido que codifica una proteína de esta invención puede usarse para la producción de una proteína COASTER recombinante para servir, junto con una proteína receptora nuclear recombinante, en un ensayo para la unión de la proteína con el receptor nuclear sensible.

La expresión receptor nuclear sensible se usa en esta descripción para referirse a cualquier receptor nuclear que puede activarse por una proteína COASTER. Preferiblemente, el receptor nuclear en el sistema del ensayo es un receptor de esteroides o, más preferiblemente, un receptor de progestágenos, un receptor de estrógenos o un receptor de glucocorticoides. Un método de modulación de la actividad transcripcional promovida por un receptor nuclear sensible y un coactivador de acuerdo con esta invención puede ser un ensayo *in vitro* para la determinación del grado de modulación, y comprende cuantificar el grado de modulación. Las condiciones para la cuantificación del grado de modulación o de unión pueden comprender un ensayo de doble híbrido en levaduras, como se describe en la presente memoria, o un ensayo de doble híbrido análogo en un sistema de células de mamíferos. Como alternativa, la eficacia de unión puede medirse usando proteínas recombinantes aisladas de COASTER y el receptor nuclear que estén convenientemente marcados para permitir la detección de la eficacia de unión. El marcaje de las proteínas puede implicar moléculas radiactivas o fluorescentes o moléculas con propiedades enzimáticas que estén directa o indirectamente unidas a las proteínas recombinantes. Como alternativa, la eficacia de unión puede medirse mediante un sistema biodetector basado en afinidad tal como el BIACORE. Como alternativa, la eficacia de unión puede medirse mediante ensayos de complementación *in vivo* o *in vitro* que usan la interacción de la proteína COASTER y del receptor nuclear para dirigir la complementación de una proteína funcional, tal como β -galactosidasa o dihidrofolato reductasa, mediante la cual la función de la proteína complementada se usa como una medida para la eficacia de unión.

El método de modulación de la actividad transcripcional promovida por un receptor nuclear sensible y un coactivador de acuerdo con esta invención también puede ser un tratamiento del cuerpo humano o animal que comprenda la administración del agente al ser humano o animal. Dicho agente puede ser un vector que inserte el gen para el coactivador en células del organismo, permitiendo de este modo la producción de más coactivador para la activación del receptor nuclear, o puede ser un vector que inserte un gen para un coactivador defectuoso en células del organismo, permitiendo de este modo la producción de una proteína de acuerdo con la invención que interfiera con la acción de la proteína COASTER endógena, dando como resultado una acción disminuida del receptor nuclear. El agente también puede ser un compuesto químico que influya en la interacción entre el coactivador y el receptor nuclear sensible. Dicho compuesto puede obtenerse mediante una exploración de rutina en un ensayo para la unión entre el coactivador y el receptor nuclear sensible, o en un ensayo para la cuantificación de la modulación de la acción del receptor nuclear sensible. Por lo tanto, un aspecto adicional de la invención es que el ensayo o método de acuerdo con la invención puede usarse para seleccionar compuestos que modulen la unión entre una proteína COASTER y un receptor nuclear sensible, o que modulen la actividad de un complejo receptor nuclear-COASTER. En dicho ensayo, se ensayan compuestos para determinar su capacidad para inducir o alterar una interacción entre COASTER y un receptor nuclear sensible o, como alternativa, sobre su capacidad para modular la transcripción mediante un receptor nuclear sensible y en presencia de COASTER y de un ADN indicador. Un ADN indicador consiste en un promotor que es sensible al receptor nuclear bajo estudio y un gen indicador cuya actividad transcripcional puede controlarse. El promotor puede ser un promotor natural aislado de un gen respuesta de un receptor nuclear o un casete sintético de (múltiples copias de) un elemento de respuesta de un receptor nuclear consenso fusionado a un promotor de núcleo basal (por ejemplo, una caja TATA). Los genes indicadores adecuados incluyen genes que codifican enzimas cuya actividad puede controlarse mediante un ensayo enzimático (por ejemplo, luciferasa o β -galactosidasa), genes que codifican proteínas con propiedades fluorescentes (por ejemplo, proteína verde fluorescente) o genes cuyos niveles de transcritos pueden controlarse.

El último ensayo es particularmente ventajoso porque COASTER potencia la acción de un promotor que contiene un elemento de respuesta del receptor de estrógenos (ERE) en presencia del receptor de estrógenos alfa y de un agonista parcial, permitiendo de este modo la identificación de nuevos agonistas parciales. Conocidos ensayos basados en ERE no son capaces de reconocer todos los agonistas parciales (por ejemplo, raloxifeno).

La invención también proporciona una composición farmacéutica como una medicina para el método mencionado de modulación de la actividad transcripcional promovida por un receptor nuclear sensible y un coactivador en un sistema, que comprende la administración de un agente a un cuerpo humano o animal, interfiriendo el agente con, o potenciando, la función del coactivador de esta invención. Esta invención también proporciona una composición farmacéutica para usar en lo mencionado, comprendiendo la composición farmacéutica el agente que interfiere con, o que potencia, la función del coactivador. Como se ha explicado anteriormente, un compuesto identificado con los ensayos proporcionados por la invención puede interferir con, o potenciar, la función como coactivador de la proteína de la invención. Como alternativa, puede formularse para uso medicinal un vector que comprenda el gen para una proteína de acuerdo con la invención. Es bien conocido en la técnica un método de preparación de una medicina por mezcla del agente con uno o más coadyuvantes farmacéuticamente aceptables, tales como los que se describen en la bibliografía de referencia Gennaro *et al.*, *Remington's Pharmaceutical Sciences*, (18ª ed., Mack publishing Company, 1990, véase especialmente la Parte 8: Pharmaceutical Preparations and Their Manufacture). Se describen coadyuvantes adecuados, por ejemplo, en Handbook of Pharmaceutical Excipients (2ª Edición, Editores A. Wade y P. J. Weller; American Pharmaceutical Association; Washington; The Pharmaceutical Press; Londres, 1994). La mezcla del agente y del coadyuvante farmacéuticamente aceptable puede comprimirse en unidades de dosificación sólidas, tales como píldoras y comprimidos, o procesarse en cápsulas o supositorios. Por medio de líquidos farmacéuticamente adecuados, el agente también puede aplicarse como una preparación para inyección en forma de una solución, suspensión, emulsión o como una pulverización, por ejemplo, una pulverización nasal. Para preparar unidades de dosificación, por ejemplo, comprimidos, se contempla el uso de aditivos convencionales tales como cargas, colorantes, aglutinantes poliméricos y similares. Dichas composiciones farmacéuticas pueden usarse en un tratamiento relacionado con un receptor nuclear activado por COASTER, más específicamente, contra el desarrollo de tumores estimulados por hormonas, para la anticoncepción masculina o femenina, para el tratamiento de dolencias menopáusicas y trastornos postmenopáusicos en mujeres, para el tratamiento con fármaco anabolizantes, para enfermedades cardíacas y para el tratamiento del envejecimiento debido a una actividad hormonal reducida. Puesto que la proteína COASTER presenta un perfil de expresión específico de tejido, dichas medicinas pueden contribuir a tratamientos específicos de tejido. Además, la presencia o la modulación de COASTER pueden alterar la farmacología de ligandos de receptores nucleares. Por ejemplo, habitualmente el raloxifeno es un antagonista, pero cuando se expresa claramente COASTER en la célula, el raloxifeno presenta propiedades agonistas. También se ponen a disposición por esta invención polinucleótidos marcados adecuadamente que tienen una secuencia que codifica proteínas para la invención o fragmentos de las mismas. Estos pueden usarse en métodos de diagnóstico mediante hibridación cuantitativa rigurosa de ARN para identificar enfermedades relacionadas con una expresión anormal de ARN de COASTER en un organismo. Estos métodos pueden implicar, ventajosamente, el uso de dichos polinucleótidos marcados, que tienen una longitud de al menos 500 nucleótidos, para realizar una identificación precisa del ARN con hibridación. Sin embargo, también puede preferirse una longitud de entre 20 y 50 nucleótidos, siendo más práctica y todavía eficaz para este fin. Con fines de diagnóstico, también puede detectarse la expresión de ARNm de COASTER mediante una reacción en cadena de nucleótidos, tal como una reacción en cadena de la polimerasa con transcripción inversa (RT-PCR). Para este fin, se sintetiza ADNc a partir de ARN aislado de un espécimen de muestra usando una enzima transcriptasa inversa. El ADNc se usa posteriormente como un molde en una reacción en cadena de la polimerasa en presencia de un cebador oligonucleotídico directo e inverso de COASTER. La cantidad de producto de PCR amplificado, que puede determinarse mediante marcaje isotópico o no isotópico junto con métodos de detección apropiados o mediante métodos espectrofotométricos o colorimétricos directos, refleja el nivel de expresión relativo del ARNm de COASTER.

El marcaje puede realizarse de diversas formas. En el caso de marcaje radiactivo, la hibridación de polinucleótidos puede medirse mediante la cuantificación del número de desintegraciones por segundo del híbrido radiomarcado. Como alternativa, los fragmentos radiomarcados hibridados pueden visualizarse mediante autorradiografía o métodos similares. Los métodos de marcaje no isotópicos con frecuencia hacen uso de nucleótidos biotinilados que se incorporan durante el marcaje del fragmento. Posteriormente, el fragmento hibridado con éxito puede detectarse en virtud de la biotina incorporada. El resto de biotina puede detectarse usando una molécula de avidina conjugada con una molécula con propiedades fluorescentes o enzimáticas. La cantidad de actividad fluorescente o enzimática se usa como una medida para la eficacia de hibridación. El método de diagnóstico será particularmente útil para enfermedades que implican desarrollo tumoral. Es sorprendente que el gen de COASTER está muy expresado en líneas celulares tumorales. El método de diagnóstico puede realizarse mediante exploración o toma de biopsias del organismo que se va a diagnosticar, pero el método puede limitarse estrictamente a un método que no se practique en el cuerpo humano o animal. En esa situación, el método comprende el ensayo *in vitro* del material biológico disponible para hibridación. En vista de la sensibilidad de los receptores nucleares para COASTER, el método se usa preferiblemente cuando la enfermedad que se sospecha está relacionada con un mal funcionamiento de un receptor de esteroides o, más preferiblemente, con un receptor de progestágenos, un receptor de estrógenos o un receptor de glucocorticoides.

Las proteínas para la invención también pueden usarse para producir anticuerpos dirigidos contra estas proteínas, pudiendo usarse los anticuerpos en métodos de diagnóstico de trastornos que implican cambios en el proceso de coactivación descrito en esta invención. Pueden encontrarse métodos para la producción de anticuerpos monoclonales y policlonales en un manual de laboratorio habitual tal como Current Protocols in Molecular Biology (F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, K. Struhl eds., John Wiley y Sons Inc.). Puede llevarse a cabo la inmunización usando preparaciones en bruto o purificadas de proteína COASTER o mediante vacunación de ADN usando un vector de expresión de mamíferos para COASTER que se usa directamente para inmunizar el organismo donante de anticuerpos.

Pueden obtenerse polinucleótidos que codifican COASTER incorporados artificialmente en células eucariotas, bacterias, plásmidos o vectores y proteína COASTER por amplificación mediante una reacción en cadena de la polimerasa sobre un molde de ADNc que se haya obtenido mediante la transcripción inversa de ARN de testículo humano. La reacción en cadena de la polimerasa debería hacer uso de un cebador directo que comprende el codón de inicio del ADNc de COASTER (por ejemplo, ATGGGAGACCCGGGGTTCGGA; SEC ID 13) y un cebador inverso que comprende el codón de terminación del ADNc de COASTER (por ejemplo, TTACATTTCCTCAAGACTTC; SEC ID 14). La clonación del ADNc obtenido de este modo en un vector de expresión adecuado permitirá la producción de la proteína COASTER. Mediante delecciones e inserciones de nucleótidos parcialmente idénticos pueden obtenerse proteínas COASTER artificiales con métodos bien conocidos por los especialistas, por ejemplo, como se describe en Sambrook (en la obra citada).

La proteína COASTER puede introducirse en un organismo mediante la transfección de un vector de expresión que contiene el ADNc de COASTER. El vector debería contener un promotor que sea capaz de dirigir la expresión génica en el organismo transfectado, y el ADNc de COASTER cadena abajo de este promotor. Además, el vector debería contener señales para un inicio y una terminación adecuadas de la transcripción y de la traducción, y señales para un procesamiento adecuado del transcrito. Para la introducción de COASTER en células de mamífero son adecuados promotores virales tales como el Virus de los Simios 40, de la repetición terminal larga del Sarcoma de Rous o los promotores tempranos inmediatos de Citomegalovirus, para dirigir la transcripción del ADNc de COASTER. Los ejemplos de vectores plasmídicos que contienen promotores virales incluyen pCDNA3.1, pRSV y pNGV1. Además de la introducción de COASTER en células de mamífero con la ayuda de plásmidos, COASTER puede introducirse en células de mamífero mediante transducción usando vehículos virales para el suministro y expresión de COASTER. Los virus adecuados incluyen retrovirus, adenovirus y baculovirus. Para la introducción de COASTER en células bacterianas, son adecuados vectores que contienen promotores bacterianos tales como el promotor de T7. Los ejemplos de vectores de expresión bacterianos incluyen pRSET y pET. Para la introducción de COASTER en levaduras, los vectores adecuados incluyen pYES para la introducción en *Saccharomyces cerevisiae* y pPIC para la introducción en *Pichia pastoris*.

Está disponible un método sencillo para seleccionar un receptor nuclear sensible para el ensayo. Por ejemplo, puede ensayarse si un receptor nuclear se coactiva mediante una proteína COASTER en un ensayo de transfección transitoria. Cuando existe coactivación, el receptor nuclear es sensible. La acción de un receptor nuclear es la inducción de la transcripción de un gen unido cadena abajo a un promotor. Con un ensayo de doble híbrido, por ejemplo, puede ensayarse si el receptor nuclear interacciona con la proteína COASTER. Asimismo, puede ensayarse de forma rutinaria la equivalencia funcional con la proteína COASTER de una proteína artificial en un ensayo como se describe en los ejemplos.

La siguiente descripción de variantes y uso de la invención pretende posibilitar adicionalmente la invención y el uso de la misma.

La coactivación del ER α unido a un ligando modulador selectivo del receptor de estrógenos (SERM) depende de la región AF1

La Figura 4 muestra que COASTER es un coactivador para el ER α unido a ligando de 4OH-tamoxifeno pero no para el ER β unido a ligando de 4OH-tamoxifeno. Estos resultados están muy en concordancia con observaciones anteriores de que el SERM 4OH-tamoxifeno produce una señal mediante la función AF1 del ER α . La función AF1 reside en el dominio AB del ER α y no está ni estructuralmente ni funcionalmente conservada en ER β (McInerney *et al.*, 1998, Endocrinology 139, 4513-4522). Para confirmar esta idea, se ensayó el potencial de coactivación de COASTER sobre dos receptores quimera. La quimera ER β / α contiene el dominio AB de ER β fusionado con los dominios CDEF de ER α , como se ha descrito anteriormente (McInerney *et al.*, 1998, Endocrinology 139, 4513-4522). En experimentos de transfección, el receptor ER β / α permite la coactivación mediante COASTER en presencia de 17 β -estradiol pero no en presencia de los SERM 4OH-tamoxifeno y raloxifeno (Figura 8). Esta observación indica que la coactivación mediante COASTER del receptor unido a un ligando SERM es distinta de la coactivación del receptor unido a ligando de 17 β -estradiol. Conforme a la observación anterior de que la señalización mediante 4OH-tamoxifeno requiere el AF1 de ER α , los resultados demuestran que el 4OH-tamoxifeno no presenta agonismo sobre la quimera ER β / α . Además, los datos demuestran que el aumento de los niveles de COASTER en la célula no pueden superar este bloqueo en la actividad, indicando que la coactivación del receptor unido a un ligando SERM requiere la AF1 de ER α . La situación en la quimera ER α / β , que contiene la AF1 de ER α fusionada con los dominios CDEF de ER β , es más compleja. Se ha demostrado anteriormente que la quimera ER α / β muestra agonismo de 4OH-tamoxifeno y una actividad en presencia de 17 β -estradiol que está reducida en comparación con los receptores ER α o ER β de tipo silvestre (McInerney *et al.*, 1998, Endocrinology 139, 4513-4522). Los datos confirman y amplían estos resultados y demuestran que, además del 4OH-tamoxifeno y del 17 β -estradiol, el raloxifeno también es capaz de producir una señal a través de la quimera ER α / β . Sin embargo, la quimera ER α / β es insensible a la coactivación por COASTER. Los resultados combinados indican que la quimera ER α / β adopta una conformación única y sugieren que la coactivación, como se observa en el ER β de tipo silvestre unido a ligando de 17 β -estradiol, depende tanto de la función de AF2 como de las secuencias en el extremo N-terminal de ER β .

COASTER identifica nuevos SERM

Los resultados que se presentan en la Figura 4 y 5 muestran que en ausencia de COASTER, el 4OH-tamoxifeno puede funcionar como un agonista sobre ER α , mientras que el raloxifeno no puede hacerlo. La elevación de los niveles de COASTER en la célula permite la identificación de raloxifeno como un ligando con potencial agonista. Se planteó

si COASTER podría permitir la identificación de nuevos SERM potenciales en una combinación de compuestos que se hubieran identificado previamente como antiestrógenos. Para este fin, se ensayó la actividad de los compuestos ORG 43143, ORG 39660 y ORG 39669, que se identificaron como nuevos antiestrógenos en un ensayo de exploración de alto rendimiento (Method in Dechering *et al.*, 2000, Current Medicinal Chemistry 7, 561-576). La Figura 9 muestra que estos compuestos se comportan todos como antiestrógenos. En particular, el compuesto ORG 39669 muestra una actividad parcial en el ensayo de antiestrógenos. Cuando se ensayaba en un ensayo de coactivación de COASTER, este grupo de antiestrógenos se clasificaba en tres clases funcionales diferentes (Figura 10). La primera clase comprende 4OH-tamoxifeno y ORG 39669 y representa un grupo de compuestos que muestran cierto agonismo sobre ER α y un agonismo potenciado en presencia de COASTER. La segunda clase es un grupo de antiestrógenos que están silentes en la transcripción mediada por ER α de un indicador de ERE en ausencia de COASTER. Sin embargo, cuando los niveles de COASTER están elevados, estos compuestos se vuelven agonistas. Los compuestos raloxifeno y el nuevo antiestrógeno ORG 43143 son representativos de esta segunda clase. La última clase de antiestrógenos incluye ICI 164.384 y ORG 39660 y representa un grupo de compuestos que no muestran agonismo sobre un indicador de ERE y permanecen silentes en presencia de COASTER.

Distribución tisular de COASTER

La distribución tisular de ARNm de COASTER se investigó mediante análisis de transferencia de Northern y RT-PCR. Se hibridaron transferencias de Northern de múltiples tejidos humanos con un fragmento de ADNc de COASTER de 794 pb. Casi todos los tejidos examinados muestran hibridación con una especie de ARNm de aproximadamente 5 kb de tamaño, con mayor nivel de señal en testículo, ovario, corazón, placenta y músculo esquelético y señales de inferiores a nulas en los otros tejidos (Figura 6). Además, se observa una señal de hibridación muy intensa con un transcrito de 4 kb que sólo está presente en el testículo. El tamaño del transcrito de 5 kb está en buena concordancia con el tamaño del ADNc que se ha aislado. El transcrito de 4 kb puede representar una variante de corte y empalme. El patrón de expresión específico del testículo de este transcrito es sorprendente y sugiere un papel funcional en la fisiología testicular. También se investigó la expresión de ARNm de COASTER mediante análisis de RT-PCR en diferentes líneas celulares humanas procedentes de tumores y en muestras de ARN de tejidos humanos y de *Macaca fasciculata* (mono *Cynomolgus*). Los resultados semicuantitativos de la RT-PCR se muestran en las Figuras 7 y 11. Los resultados muestran una elevada expresión en las líneas celulares de carcinoma epitelial de mama MCF-7 y T-47 y en la línea de células epiteliales vaginales SW954. Se observó una expresión moderada en las líneas celulares de osteosarcoma U-2 OS y HOS, en la línea celular de carcinoma epitelial endometrial Ishikawa, en la línea de células endoteliales vasculares VE103ER α y en la línea celular endotelial HS760T. No se detectó expresión en la línea celular de osteosarcoma MG63 ni en la línea celular de carcinoma epitelial endometrial ECC-1. Se ha descrito que las líneas celulares que muestran una elevada expresión de COASTER (MCF-7, T-47D y SW954) contienen elevados niveles de ER α (Dechering *et al.*, Curr. Med. Chem. 7, 561-576), sugiriendo que COASTER puede contribuir al desarrollo de células tumorales mediado por ER α . Los resultados de RT-PCR sobre ARN procedente de tejidos se presentan en la Figura 11. El método de RT-PCR confirma la elevada expresión en testículo que reveló la transferencia de Northern (Figura 6). Además, se observa una elevada expresión en tejido óseo. Esto indica que los ligandos que permiten la interacción receptor nuclear-COASTER podrían influir en el metabolismo del hueso.

Los ejemplos y figuras adicionales sirven para ilustrar y clarificar o especificar la invención. La descripción se ilustra con las siguientes figuras:

Figura 1. Ejemplo de una secuencia de nucleótidos y de aminoácidos de COASTER. Esta secuencia de ADNc de COASTER tiene una longitud de 4999 pares de bases con una fase de lectura abierta de 3039 pares de bases (1013 aminoácidos). Una secuencia previamente publicada del ADNc que codifica KIAA0576 (número de acceso AB011148) con una homología casi total (99%) muestra una inserción de 144 nucleótidos en la posición 2832 que da como resultado una fase de lectura abierta de 1061 aminoácidos. Se muestran en negrita los motivos de dedos de cinc y se subraya un motivo LXXLL. La secuencia de aminoácidos de una señal de localización nuclear consenso se muestra en cursiva y subrayada. El codón de inicio ATG y los codones de terminación de la traducción en la fase de lectura abierta se muestran en negrita y cursiva.

Figura 2. La secuencia de ADNc de COASTER codifica una proteína de aproximadamente 115 kiloDaltons. El molde pCDNA3.1HISC.COASTER que contiene la fase de lectura abierta completa de COASTER se transcribió y se tradujo *in vitro* en presencia de [³⁵S]-metionina. Como control, se transcribió y se tradujo pBK-SRC-1 (Onate *et al.*, Science 270, 1354-1357) en paralelo. Se analizaron las proteínas marcadas mediante SDS-PAGE. La Figura muestra una autorradiografía del patrón de electroforesis. Carril 1: producto de traducción de pBK-SRC-1. Carril 2: producto de traducción de pCDNA3.1HISC.COASTER. La posición de la proteína COASTER se indica con una flecha.

Figura 3. Interacción dependiente de hormonas entre ER β y COASTER. Se realizó un ensayo de doble híbrido en levaduras en la cepa hospedadora de levaduras AH109, que se transformó con plásmido pGBT9.ER β y con un plásmido que codifica una fusión entre un dominio de activación GAL4 y los aminoácidos 1 a 234 de COASTER. La Figura muestra la actividad β -galactosidasa de las células de levadura transformadas en respuesta a 17 β -estradiol 10⁻⁸ M (barra sombreada) y raloxifeno 10⁻⁵ M (barra negra). La actividad β -galactosidasa de fondo de las células de levadura transformadas se indica con una barra blanca.

Figura 4. Coactivación diferencial de ER α y ER β en presencia de 4OH-tamoxifeno y raloxifeno. Se transfectaron de forma transitoria células U-2 OS con p4ERE-TATA-LUC y pNGV1ER α o pNGV1ER β como se indica. Cuando se indica, las células se cotransfectaron con las construcciones de expresión de COASTER pCDNA3.1HISA.COASTER Δ o pCDNA3.1HISC.COASTER, o con la construcción de expresión pCDNA3.1HISC.GRIP1. La figura muestra la actividad luciferasa (a. l.) después de la incubación de las células sin hormona (barras blancas), con raloxifeno 10⁻⁷ M (barras sombreadas) o con 4OH-tamoxifeno 10⁻⁷ M (barras negras). Las barras de error indican las desviaciones típicas en un experimento triplicado.

Figura 5. Activación dependiente de la dosis de un elemento de respuesta a estrógenos en presencia de ER α o ER β junto con COASTER. Se transfectaron de forma transitoria células U-2 OS con p4ERE-TATA-LUC y pNGV1ER α . Los experimentos de transfección indicados con marcadores negros se suplementaron con la construcción de expresión de COASTER pCDNA3.1HISC.COASTER. La figura muestra la actividad luciferasa (a. l.) después de la incubación de las células con una concentración creciente de raloxifeno (indicada con cuadrados) o 4OH-tamoxifeno (indicada con triángulos). Las barras de error indican las desviaciones típicas en un experimento triplicado.

Figura 6. Distribución tisular de transcritos de COASTER. Se hibridaron transferencias de Northern de múltiples tejidos humanos con una sonda de ADNc de COASTER marcada con [³²P]. La figura muestra una autorradiografía de las transferencias. La sonda de COASTER hibrida con dos especies de ARNm de aproximadamente 5 kb y 4 kb de tamaño, señaladas con flechas. Los tejidos que se representan en las transferencias incluyen bazo (a), timo (b), próstata (c), testículo (d), ovario (e), intestino delgado (f), colon (g), leucocitos de sangre periférica (h), corazón (i), cerebro (j), placenta (k), pulmón (l), hígado (m), músculo esquelético (n), riñón (o) y páncreas (p). El patrón de migración de un patrón de tamaños se indica en el lado izquierdo de la figura (tamaño en kilobases).

Figura 7. Expresión de COASTER en líneas celulares tumorales. Los niveles de expresión relativos de COASTER se determinaron mediante RT-PCR. La figura muestra la intensidad relativa de los productos de PCR medidos mediante exploración de imágenes de electroforesis en gel de agarosa de los productos de PCR. Las líneas celulares que se representan incluyen: MCF-7 (a), T-47D (b), U-2 OS (c), MG63 (d), HOS (e), ECC-1 (f), Ishikawa (g), VE103 (h), HS760T (i) y SW 954 (j).

Figura 8. Expresión de luciferasa cuantificada con recuentos de luz por segundo (cps) en células transfectadas con un receptor quimérico ER β/α y COASTER en respuesta a la adición de estradiol (E2), raloxifeno o 4-hidroxi-tamoxifeno (4OHT).

Figura 9. Porcentaje de actividad estrogénica en respuesta a la adición de una dosis patrón de estradiol en presencia de diversas concentraciones, expresadas como logaritmo de la concentración, de raloxifeno, 4-hidroxi-tamoxifeno (4OHT), ICI 164.384, ORG 43143, ORG 39660 y ORG 39669, respectivamente.

Figura 10. Efecto de compuestos en un ensayo de coactivación de COASTER.

Figura 11. Expresión de COASTER en tejidos humanos y de *M. fasciculata*. Los niveles de expresión relativa de COASTER se determinaron mediante RT-PCR. La figura muestra la intensidad relativa de los productos de PCR medidos mediante un método de PCR cuantitativa a tiempo real Taqman. Los tejidos que sirvieron como la fuente para el material de ARN se indican en la parte inferior de la figura. Todos los tejidos eran de origen humano, excepto por el tejido óseo que procedía de *M. fasciculata*.

Ejemplos

Identificación de un nuevo coactivador para receptores de esteroides

Resultados y discusión

Exploración de doble híbrido en levaduras

Para la identificación de COASTER como proteína que interacciona con ER β de una forma dependiente de hormonas, se realizó una exploración de doble híbrido en levaduras usando ER β como cebo. El sistema de doble híbrido consistía en una cepa de levaduras (AH109) que expresaba los genes indicadores *HIS3*, *MEL1*, *ADE2* y *lacZ* de un promotor sensible a GAL. ER β se expresaba en esta cepa en una fusión con un dominio de unión al ADN de GAL4. En este sistema, ER β solo no es capaz de activar los genes indicadores en presencia de 17 β -estradiol (no se muestran los resultados). La cepa hospedadora que expresaba ER β se transformó con una genoteca de ADNc de osteosarcoma humano. Esta genoteca expresa ADNc como una fusión traduccional con un dominio transactivador GAL4. La interacción entre ER β y una proteína codificada por una genoteca permite la activación de los genes indicadores y el crecimiento en medios que carecen de histidina y/o adenina. Se exploraron un total de 3,5 x 10⁶ colonias. Las células de levadura transformadas se seleccionaron por su crecimiento en medios que carecían de histidina y en presencia de 17 β -estradiol. Se seleccionaron 560 colonias *HIS*⁺ y se exploraron mediante réplica en placa sobre su fenotipo *ADE*⁺ en ausencia o presencia de 17 β -estradiol. Se seleccionaron 27 transformantes de levaduras que mostraban un fenotipo *HIS*⁺ y *ADE*⁺ dependiente de hormonas. El ADN plasmídico procedente de la genoteca se aisló de estos 27 clones y se reintrodujo en la cepa de levadura hospedadora AH109.pGBT9.ER β para verificar que los fenotipos observados eran dependientes del

ADN transformado. En un ensayo de transactivación de β -galactosidasa (*lacZ*) en levaduras, 21 de los 27 positivos que se volvieron a ensayar mostraron un fenotipo *lacZ*⁺ inducible por hormonas. El análisis de la secuencia de estos positivos identificó dos coactivadores bien conocidos para receptores de esteroides, UBC9 (enzima de conjugación de ubiquitina) y SUG1. Además, se identificó una secuencia de ADNc que se predijo que codificaba una proteína nuclear con varios motivos de dedos de cinc y con un dominio LXXLL, que aparentemente estaban implicados en la interacción entre receptores de esteroides y coactivadores. Este ADNc se seleccionó para un estudio adicional y se demostró que codificaba un coactivador funcional para receptores de esteroides. Por lo tanto, se ha denominado COASTER a este nuevo coactivador.

COASTER: un nuevo coactivador para receptores de esteroides

Estructura primaria de COASTER

Se secuenció el inserto del plásmido pACT2.COASTER Δ que se aisló en la exploración de doble híbrido y esta secuencia se comparó con las bases de datos Incyte Lifeseq Gold y Genbank usando el algoritmo BLAST. Se identificaron un total de 270 secuencias de EST Incyte. Se obtuvo el clon 2905757, que contenía el extremo 3' del ADNc de COASTER, de Incyte Genomics (Palo Alto, Estados Unidos) y se secuenció en su totalidad. Además, la búsqueda en la base de datos Genbank reveló que la secuencia de COASTER comparte una homología casi total (99%) con el gen que codifica KIAA0576, una supuesta proteína de función desconocida (Nagase *et al.*, DNA Res. 5, 31-39). La secuencia de KIAA0576 representa una variante de corte y empalme de COASTER y contiene una inserción en la fase de lectura abierta de 144 pares de bases. Se empleó RACE PCR para identificar el extremo 5' de la secuencia de ADNc de COASTER. Se identificaron dos fragmentos diferentes que se diferenciaban en sus regiones 5' no traducidas. La Figura 1 muestra una visión de conjunto de la secuencia completa de ADNc de COASTER. El ADNc que se corresponde con el fragmento de RACE más largo tiene una longitud de 4099 nucleótidos y contiene una fase de lectura abierta de 3039 nucleótidos (1013 aminoácidos). El KIAA0576 muestra una inserción de 144 nucleótidos en la posición 2832 que da como resultado una fase de lectura abierta de 1061 aminoácidos. Tanto para COASTER como para KIAA0576, la secuencia de aminoácidos deducida a partir del ADNc predice una proteína que contiene 9 dominios de dedos de cinc (de tipo C₂H₂), un sitio de localización nuclear y un motivo de interacción con receptores nucleares LXXLL.

El ADNc de COASTER codifica una proteína

Para ensayar si el ADNc de COASTER codificaba de hecho una proteína, se transcribió y se tradujo *in vitro* el plásmido pCDNA3.1HISC.COASTER que contiene la fase de lectura abierta completa de COASTER (aminoácidos 1-1013). La Figura 2 muestra que una proteína de aproximadamente 115 kiloDaltons está codificada por el ADNc de COASTER. El tamaño observado de la proteína está en perfecta concordancia con la longitud de la secuencia de ADNc que se presenta en la Figura 1. Los resultados muestran además que el ADNc de COASTER codifica una proteína, que es diferente de la proteína codificada por pBK-SRC-1, y que contiene la fase de lectura abierta para el coactivador SRC-1 conocido.

COASTER interacciona con ER β en levaduras

El plásmido pACT2.COASTER Δ se aisló originariamente en la exploración de doble híbrido en levaduras. Para verificar que la proteína codificada por este plásmido interacciona de hecho con ER β en levaduras, se reintrodujo el plásmido pACT2.COASTER Δ en la cepa receptora AH109.pGBT9.ER β . La Figura 3 muestra los resultados de un ensayo de doble híbrido posterior. Los resultados muestran que la adición de 17 β -estradiol al medio de cultivo da como resultado una inducción dependiente de la dosis de la actividad indicadora de β -galactosidasa. Por el contrario, el antiestrógeno raloxifeno no es capaz de inducir la actividad del indicador. Esta incapacidad puede estar relacionada con la resistencia de células de levaduras a antiestrógenos, como se ha publicado anteriormente (Lyttle *et al.*, J Steroid Biochem. Biol. 42, 677-685; IH, resultados no publicados). Como alternativa, puede indicar que la conformación inducida por antiestrógenos de ER β no permite una interacción con COASTER. La transformación de la cepa de levadura con el plásmido de expresión de ER β o con el plásmido de expresión de COASTER en solitario no dio como resultado un fenotipo *lacZ*⁺ inducible por 17 β -estradiol (no se muestran los datos). En su conjunto, los resultados indican que ER β y COASTER interacción de una forma dependiente de 17 β -estradiol en levaduras. La interacción es independiente del motivo LXXLL de COASTER, puesto que este motivo no está codificado por el plásmido pACT2.COASTER Δ que se usó en el ensayo de doble híbrido en levaduras.

COASTER es un coactivador funcional en células de mamíferos

Para ensayar el papel funcional de COASTER en un entorno de mamíferos, se clonó el ADNc de longitud completa de COASTER (aminoácidos 1-1013) en un vector de expresión de células de mamíferos. El papel funcional de COASTER se analizó mediante la cotransfección de células U-2 OS con el plásmido de expresión de COASTER junto con un plásmido que expresaba un receptor de esteroides y un plásmido que contenía un gen indicador de luciferasa bajo el control de un elemento de respuesta a hormonas. Después de la transfección, las células se estimularon con la hormona apropiada para el receptor bajo estudio. Los resultados se presentan en la Tabla 1.

TABLA 1

Células de ensayo		Sin hormona	17 β -estradiol 10 ⁻⁸ M	ICI 164.384 10 ⁻⁸ M
1	ERβ	150 $\pm$ 14	500 $\pm$ 109	134 $\pm$ 28
	ERβ + COASTER	133 $\pm$ 16	8200 $\pm$ 819	101 $\pm$ 6
2	ERα	180 $\pm$ 22	9505 $\pm$ 2122	174 $\pm$ 18
	ERα + COASTER	134 $\pm$ 18	15400 $\pm$ 1179	275 $\pm$ 55
			Org 2058 10⁻⁸ M	
3	PR	62 $\pm$ 22	822 $\pm$ 112	
	PR + COASTER	81 $\pm$ 48	3229 $\pm$ 683	
			Dexametasona 10⁻⁸ M	
4	GR	44 $\pm$ 3	1232 $\pm$ 82	
	GR + COASTER	103 $\pm$ 33	35245 $\pm$ 8679	

Leyenda de la Tabla 1: COASTER es un coactivador para receptores de esteroides. Se transfectaron de forma transitoria células U-2 OS con pNGV1.ER β y p4ERE-TATA-LUC (1), pNGV1.ER α y p4ERE-TATA-LUC (2), pKCRE.PR y pMMTV-LUC (3) o pNGV1.GR y pMMTV-LUC (4). Donde se indica, se añadió la construcción de expresión de COASTER pCDNA3.1HISC.COASTER a la mezcla de transfección. Después de la transfección, las células se incubaron con las hormonas que se indican en la fila superior. La tabla muestra actividades de luciferasa $\pm$ desviaciones típicas a partir de un experimento triplicado. Org 2058 = 16 α -etil-21-hidroxi-19-norpregn-4-eno-3,20-diona.

Como puede observarse en la Tabla 1, COASTER es un coactivador para el receptor de progesterona (PR), para el receptor de glucocorticoides (GR) y para los receptores de estrógenos α y β (ER α y ER β). En todos los casos, la actividad del receptor unido a ligando con un agonista se potencia adicionalmente por la coexpresión de COASTER. Cuando ER α o ER β están unidos a ligando con el antagonista ICI164.384, la coexpresión de COASTER no potencia la actividad transcripcional, indicando que el antagonista induce una conformación que no permite a COASTER funcionar como un coactivador. La actividad del gen indicador no se indujo cuando COASTER se transfectó en ausencia de un receptor de esteroides (no se muestran los datos), indicando que tanto el receptor de esteroides como COASTER son necesarios para la actividad potenciada.

COASTER identifica agonistas parciales para el receptor de estrógenos α

El potencial coactivador de COASTER sobre ER α y ER β se estudió adicionalmente en presencia de los antiestrógenos raloxifeno y 4OH-tamoxifeno. Para una comparación, se analizó en paralelo la actividad del coactivador GRIP1, que es un ortólogo de ratón de TIF2/SRC-2 humano. Los experimentos de transfección mostraron que en células U-2 OS, el 4OH-tamoxifeno actúa como un agonista sobre la actividad transcripcional de ER α , pero no de ER β , cuando se usa un indicador de ERE como lectura (Figura 4). La cotransfección de COASTER o GRIP1 da como resultado una actividad potenciada del ER α unido a ligando de 4OH-tamoxifeno. Curiosamente, COASTER es capaz de activar la transcripción en presencia de raloxifeno, compuesto que de otro modo no es un agonista en este sistema. Este fenómeno se observó en la transcripción mediada por ER α , pero no cuando estaba presente ER β como el receptor de estrógenos. Además, el coactivador GRIP1 no es capaz de activar la expresión del gen indicador en presencia de raloxifeno y ER α . Los datos muestran que la actividad de la forma truncada de COASTER, COASTER Δ , es idéntica a la de COASTER de longitud completa. Este indica que los aminoácidos 1 a 234 son suficientes para el potencial coactivador de COASTER sobre el ER α unido a ligando de raloxifeno o 4OH-tamoxifeno. Los efectos de COASTER sobre la transcripción mediada por ER α se estudiaron adicionalmente haciendo curvas de dosis-respuesta para los ligandos raloxifeno y 4OH-tamoxifeno en presencia y ausencia de COASTER. Los datos que se presentan en la Figura 5 confirman la observación de que el raloxifeno actúa como un agonista en la transcripción génica dirigida por ERE en presencia de ER α y COASTER. Cuando se omite COASTER, el raloxifeno no es capaz de activar el gen indicador,

incluso a concentraciones elevadas del ligando. El antiestrógeno 4OH-tamoxifeno es un agonista en este sistema. Sin embargo, la expresión de COASTER aumenta la magnitud de la respuesta transcripcional. Para ambos compuestos, la actividad del indicador en presencia de ER α y COASTER es dependiente de la concentración usada en el ensayo. Los resultados que se presentan en la presente memoria muestran que COASTER puede influir en el equilibrio agonista/antagonista de antiestrógenos de perfil mixto tales como raloxifeno y 4OH-tamoxifeno. Las publicaciones anteriores han demostrado que el raloxifeno puede actuar como un agonista a través de rutas no clásicas (no ERE) sobre ER β (Zou *et al.*, Mol Endocrinol. 13, 418-430; Paech *et al.*, Science 277, 1508-1510). Hasta donde sabemos, estos datos proporcionan las primeras observaciones de agonismo de raloxifeno sobre la ruta clásica del elemento de respuesta a estrógenos (ERE). Además, éstas son las primeras indicaciones de que el raloxifeno puede actuar como un agonista sobre ER α . Estas observaciones son importantes para el entendimiento de los mecanismos de acción de los antiestrógenos. La propiedad de COASTER de activar el ER α unido a ligando de raloxifeno diferencia a COASTER de los coactivadores, tales como GRIP1, descritos hasta la fecha. Además, el ensayo que se describe en la presente memoria puede usarse con una ventaja particular para la exploración farmacológica de nuevos ligandos de receptores de estrógenos, puesto que COASTER es capaz de identificar propiedades agonistas parciales de compuestos.

Distribución tisular de COASTER

Se investigó la distribución tisular del ARNm de COASTER mediante análisis de transferencia de Northern y RT-PCR. Se hibridaron transferencias de Northern de múltiples tejidos humanos con un fragmento de ADNc de COASTER de 794 pb. Casi todos los tejidos examinados muestran hibridación con una especie de ARNm de aproximadamente 5 kb de tamaño, con mayor nivel de señal en testículo, ovario, corazón, placenta y músculo esquelético, y señales de inferiores a inexistentes en los otros tejidos (Figura 6). Además, se observó una señal de hibridación muy intensa con un transcrito de 4 kb que sólo está presente en el testículo. El tamaño del transcrito de 5 kb está en buena concordancia con el tamaño del ADNc que se había aislado. El transcrito de 4 kb puede representar una variante de corte y empalme de la que aún se desconoce la estructura detallada. Sin embargo, el patrón de expresión específico de testículo de este transcrito es sorprendente y sugiere un papel funcional en la fisiología testicular. También se investigó la expresión de ARNm de COASTER mediante análisis de RT-PCR en diferentes líneas celulares humanas procedentes de tumores. Los resultados semicuantitativos de la RT-PCR se muestran en la Figura 7. Los resultados muestran una elevada expresión en las líneas celulares de carcinoma epitelial de mama MCF-7 y T-47 y en la línea de células epiteliales vaginales SW954. Se observó una expresión moderada en las líneas celulares de osteosarcoma U-2 OS y HOS, en la línea celular de carcinoma epitelial endometrial Ishikawa, en la línea de células endoteliales vasculares VE103ER α y en la línea celular endotelial HS760T. No se detectó expresión en la línea celular de osteosarcoma MG63 ni en la línea celular de carcinoma epitelial endometrial ECC-1. Se ha descrito que las líneas celulares que muestran una elevada expresión de COASTER (MCF-7, T-47D y SW954) contienen elevados niveles de ER α (Dechering *et al.* Curr. Med. Chem. 7, 561-576), indicando que COASTER puede contribuir al crecimiento mediando por ER α de células tumorales.

Materiales y métodos

Exploración de doble híbrido en levaduras

Medio de levaduras

Se cultivaron cepas hospedadoras de levaduras y transformantes en medio completo (YPD) o en medio Synthetic Dropout (SD) mínimo. El medio completo es una mezcla de peptona, extracto de levadura y dextrosa en proporciones óptimas para el cultivo de cepas de levaduras (véanse los protocolos de levaduras del manual de Clontech, Palo Alto, Estados Unidos, protocolo PT3024-1). El medio SD mínimo contiene dextrosa al 2%, bases nitrogenadas de levaduras al 0,67% y una mezcla específica de aminoácidos y nucleósidos (por ejemplo, SD-His-Leu-Trp es un medio que carece de histidina, leucina y triptófano). Todos estos medios se adquirieron de BIO101 (Beverly, Estados Unidos).

Exploración de doble híbrido

Se obtuvo la cepa AH109 de *Saccharomyces cerevisiae* (MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4 Δ , gal80 Δ , LYS2::GAL1_{UAS}-GAL1_{TATA}-HIS3, GAL2_{UAS}-GAL2_{TATA}-ADE2, URA3::MEL1_{TATA}-lacZ) de Clontech, Palo Alto Estados Unidos. Para expresar la proteína ER β como un cebo en la cepa de levaduras AH109, se usó plásmido pGBT9.ER β (por cortesía de W. Kruijer, Rijksuniversiteit Groningen, Holanda). Este plásmido contiene la secuencia codificante de ER β (aminoácidos 1 a 530) fusionada a una secuencia que codifica un dominio de unión al ADN de GAL4 y contiene el gen *TRP1* como un marcador de selección. La cepa de levadura AH109 se transformó con pGBT9.ER β usando un método de acetato de litio convencional siguiendo el manual de instrucciones de Clontech (Palo Alto, Estados Unidos, protocolo N° PT3247-1) y se seleccionaron transformantes en placas de SD-triptófano (SD-Trp). De esta placa se seleccionaron colonias individuales y se guardaron como clones parentales. Posteriormente, la expresión de fondo de *HIS3* de estos clones se determinó mediante siembra en las placas de selección apropiadas. Se descartaron los clones que crecieron en medios deficientes en histidina en presencia de 17 β -estradiol. Posteriormente, los clones con baja expresión de fondo de *HIS3* se transformaron con un control positivo pGRIP1 (por cortesía de M. Stallcup, Universidad del Sur de California, Estados Unidos), un plásmido que codifica el coactivador GRIP1 que se ha demostrado que interacciona con el receptor de estrógenos en levaduras (Hong *et al.*, Mol. Cell. Biol. 17, 2735-2744). Posteriormente, se seleccionó el clon parenteral (AH109.ER β), que mostraba una elevada expresión de *HIS3* en presencia de pGRIP1 y 17 β estradiol, como el clon receptor para la exploración de la genoteca. Se obtuvo de Clontech

(Clontech, Palo Alto, Estados Unidos, n° de cat.: HL4026AH) una genoteca de ADNc Matchmaker de osteosarcoma humano. Esta genoteca se construyó en vector pACT2, que contiene un gen *LEU2* como marcador de selección y que expresa los insertos clonados en fusión con un dominio de activación GAL4 (AD). El clon receptor AH109.ER β se transformó con 100 μ g de ADN plasmídico de la genoteca de ADNc de osteosarcoma. Se seleccionaron transformantes en placas SD-Trp-Leu-His que contenían 17 β -estradiol 10⁻⁸ M y se identificaron colonias HIS⁺. Las colonias HIS⁺ se volvieron a explorar mediante réplica en placa en medios SD-Trp-Leu-His-Ade en ausencia y presencia de 17 β -estradiol 10⁻⁸ M. Los transformantes de levadura que mostraron un fenotipo HIS⁺/ADE⁺ dependiente de hormonas se seleccionaron para los estudios posteriores. A partir de estos transformantes, se aisló el ADN siguiendo un protocolo para aislamiento de minipreparaciones en bruto (véanse los protocolos básicos Y1.4). Las minipreparaciones en bruto se usaron para transformar *E. coli* KC8 mediante electroporación y se seleccionaron los transformantes que contenían el plásmido procedente de la genoteca en medio deficiente en leucina. Los insertos de la genoteca se secuenciaron en un secuenciador ABI de Perkin Elmer (Norwalk, Estados Unidos). Se realizaron búsquedas de similitud de Blast contra la base de datos Incyte.

Ensayo de gen indicador quimioluminiscente para β -galactosidasa en levaduras

Se cultivó una colonia individual en medio líquido SD-Trp-Leu durante una noche a 30°C. Estos cultivos de una noche se diluyeron hasta una DO₆₀₀ de 0,1 en 90 μ l de medio SD-Trp-Leu en una placa de cultivo de 96 pocillos. Se añadieron 10 μ l de una solución madre de hormonas concentrada 10 veces y las placas se incubaron durante 5 horas a 30°C y bajo agitación vigorosa a 200 rpm. Posteriormente, se midió la actividad β -galactosidasa usando el tampón One-Step Yeast Lysis y el sistema de detección quimioluminiscente Galacton Plus de Tropix (Tropix, Inc., Avenue, Bedford, Estados Unidos). Se controló la emisión de luz usando un Victor Instrument (Perkin Elmer, Norwalk, Estados Unidos).

RACE 5' (Amplificación rápida de extremos de ADNc) de COASTER: para aislar el extremo 5' del ADNc de COASTER, se realizaron experimentos de 5' RACE-PCR usando ADNc Marathon-Ready de testículo humano (Clontech, Palo Alto, Estados Unidos, N° Cat. 7414-1). Se realizó una RACE PCR usando un kit de amplificación de ADNc Marathon (Clontech, Palo Alto, Estados Unidos) junto con un cebador específico de gen CCAGACACC CACGTGTGGCC, siguiendo las instrucciones del fabricante. Se obtuvieron dos fragmentos diferentes (de ~650 pb y ~800 pb). Los fragmentos se aislaron a partir de un gel de agarosa, se purificaron usando una columna de centrifugación Quiaquick (Qiagen, Valencia, Estados Unidos) y se clonaron usando el kit pCR2.1TOPO (Invitrogen, Carlsbad, Estados Unidos). Las colonias que contenían insertos se identificaron mediante PCR de colonias usando los cebadores específicos de gen CCAGACACCCACGTGTGGCC y AAGCCACCATGGACAGCAGGAGCAGTGGTG. Se realizó el análisis de secuencia de los clones positivos en un secuenciador ABI (Perkin Elmer, Norwalk, Estados Unidos).

Plásmidos recombinantes

Inicialmente, se aisló un plásmido que codificaba una forma truncada de COASTER (pACT2.COASTER Δ , aminoácidos 1-234) en la exploración de doble híbrido en levaduras. El inserto de este plásmido se clonó en el vector de expresión de células de mamíferos pCDNA3.1HISA (Invitrogen, Carlsbad, Estados Unidos) para dar el plásmido pCDNA3.1HISA.COASTER Δ . Se obtuvo un plásmido que contenía el extremo 3' del ADNc de COASTER de Incyte Genomics (Palo Alto, Estados Unidos). Este plásmido (Incyte clon ID 2905757) se digirió con *SacI/NotI*. El extremo 5' del COASTER (aminoácidos 1 a 332) se amplificó por PCR sobre ADNc Marathon Ready de testículo humano (Clontech, Palo Alto, Estados Unidos, n° cat. 7414-1) usando los cebadores CTAGGTACCGGACAGCAGGAGCAGTGGT GC (SEC ID 15) y GGCAGTGAGCTCCATGTGGG (SEC ID 16) (los sitios de restricción están subrayados). El producto de PCR resultante se digirió con *KpnI* y *SacI* y se ligó junto con el fragmento de restricción *SacI-NotI* en el vector pCDNA3.1HISC digerido con *KpnI-NotI* (Invitrogen, Carlsbad, Estados Unidos). El último plásmido se denominó pCDNA3.1HISC.COASTER. Para los estudios de transfección, las regiones codificantes de proteína de los ADNc para los receptores de estrógenos α y β humanos y para el receptor de glucocorticoides se insertaron en el vector de expresión de mamíferos pNGV1 (número de acceso X99274). La región codificante de proteína del receptor de progesterona se insertó en el vector de expresión de mamíferos pKCRE (Stam *et al.*, Eur. J. Pharmacol. 227, 153). El plásmido indicador pMMTV-LUC se construyó por modificación del vector pManMneoLUC (Clontech, Palo Alto, Estados Unidos). El casete de neomicina se eliminó por digestión con endonucleasas de restricción *HindIII/BamHI*, seguido de generación de extremos romos con polimerasa Klenow y ligación mediante ADN ligasa T4. El vector indicador 4ERE.TATALuc contiene un gen de luciferasa bajo el control de cuatro elementos de respuesta a estrógenos (ERE) y se obtuvo por cortesía de P. v. d. Saag (Nederlands Instituut voor Ontwikkelingsbiologie, Utrecht, Holanda). Se construyó un plásmido de expresión de mamíferos para GRIP1 mediante digestión de restricción de pGRIP1 (por cortesía de M. Stallcup, Universidad del Sur de California, Estados Unidos) con *EcoRI* y *BamHI*. El fragmento resultante se ligó en el vector pCDNA3.1HISC digerido con *EcoRI/BamHI* para dar el plásmido pCDNA3.1HISC.GRIP1.

Cultivo celular y transfección transitoria

Se obtuvieron células U-2 OS de ATTC (HTB-96) y se mantuvieron a 37°C en una atmósfera humidificada (CO₂ al 5%) como un cultivo en monocapa en medio M505 sin rojo fenol. El último medio consiste en una mezcla de medio de Eagle modificado por Dulbecco (1:1) (DMEM, Gibco BRL, Breda, Holanda) y medio de nutrientes F12 (F12 de Ham, Gibco BRL, Breda, Holanda) suplementados con carbonato sódico 2,5 mg/ml, piruvato sódico 55 μ g/ml, β -mercaptoetanol 2,3 μ g/ml, etanolamina 1,2 μ g/ml, L-glutamina 360 μ g/ml, selenito sódico 0,45 μ g/ml, estreptomina 62,5 μ g/ml y suero de ternero bovino tratado con carbón vegetal al 5% (Hyclone). Se sembraron 1·10⁵ células en

placas de cultivo de tejidos de 24 pocillos y se introdujo ADN mediante el uso de lipofectina (Gibco BRL, Breda, Holanda). Con este fin, se mezclaron ADN plasmídicos en 62,5 μ l de Optimem (Gibco BRL, Breda, Holanda) con un volumen igual de reactivo de lipofectina diluido (1,75 μ l de lipofectina en 62,5 μ l de Optimem) y se dejaron reposar a temperatura ambiente durante 30 minutos. Después de lavar las células una vez con medio M505 sin suero, se añadieron 250 μ l de Optimem junto con la mezcla de ADN-lipofectina a las células. Después de la incubación durante un periodo de 5 horas a 37°C (CO₂ al 5%) las células se lavaron una vez con medio M505 suplementado con suero de ternero bovino tratado con carbón vegetal al 5% y se añadieron hormonas al medio. Las células se incubaron durante una noche a 37°C. Al día siguiente, las células se lavaron una vez con solución salina tamponada con fosfato (PBS) y se prepararon extractos celulares por adición de 75 μ l de tampón de lisis (tampón fosfato 0,1 M a pH 7,8, Triton X-100 al 0,2%). Después de la incubación durante 5 min a temperatura ambiente, se añadió una alícuota de 30 μ l a 50 μ l de reactivo de ensayo luciferasa (Promega; Wisconsin, Estados Unidos). Se midió la emisión de luz usando un Victor Instrument (Perkin Elmer, Norwalk, Estados Unidos).

Análisis de transferencia de Northern

Se digirió el plásmido pCDNA3.1HISA.COASTER Δ con *Eco*RI y *Xba*I para generar una sonda (nucleótidos 183-928). El fragmento de ADN se marcó con [α ³²P]dCTP con perlas de marcaje de ADN Ready-To Go siguiendo las instrucciones del fabricante (Pharmacia). Se hibridaron transferencias de Northern de múltiples tejidos humanos (Clontech, Palo Alto, Estados Unidos, n° cat. 7760-1 y 7759-1) con el fragmento de ADN marcado con α ³²PdCTP (véanse los protocolos básicos R1.5). Las transferencias de Northern se lavaron usando las siguientes condiciones de lavado: SSC 2X/SDS al 0,1% a 65°C dos veces durante 30 min y SSC 1x/SDS al 0,1% a 65°C una vez durante 15 min.

Análisis de RT-PCR de la expresión de COASTER en líneas celulares

Se preparó ADNc usando 2 μ g del ARN total aislado a partir de varias líneas celulares humanas usando RNazol B (Cinna/Biotech), usando el kit SuperscriptII (BRL). Se usó una porción de ADNc para la amplificación específica por PCR de COASTER (cebadores directo TGAGGCATCTGAGTCAACA e inverso CGAGCCAGAGTTAAAGCA). Se tomaron muestras después de 20, 25, 30 y 35 ciclos de PCR. Las muestras de PCR se analizaron en geles de agarosa al 1,5%. Las imágenes en gel se cuantificaron con Imagequant 5.0 (Molecular dynamics).

Traducción in vitro

Se usó un TnT Coupled Reticulocyte Lysate Systems (Promega, Madison, Estados Unidos, n° cat. L5020) para la transcripción y traducción *in vitro* de COASTER a partir de un molde de pCDNA3.1HISC.COASTER en presencia de [³⁵S]-metionina, siguiendo las instrucciones que se proporcionan con el lisado. Se analizó un fragmento de SRC-1 codificado por el plásmido pBK-SRC-1 (Onate *et al.*, Science 270, 1354-1357) en paralelo como control. Se separaron las proteínas marcadas mediante SDS-PAGE y se visualizaron mediante autorradiografía.

REIVINDICACIONES

1. Una proteína **caracterizada** porque la proteína se selecciona del grupo constituido por:

a) una proteína que consiste en una secuencia de aminoácidos como en las SEC ID 7, 8, 11 ó 12 o la secuencia de los aminoácidos número 1 a 234 de la SEC ID 7 u 8; y

b) una proteína que consiste en una secuencia de aminoácidos que tiene una similitud de al menos el 90% con la secuencia de los aminoácidos 1 a 234 de la SEC ID 7 u 8, calculada usando la matriz de similitud blosum62 y el algoritmo de Needleman y Wunsch.

2. Uso de un polinucleótido que codifica una proteína seleccionada del grupo constituido por:

a) una proteína que tiene una secuencia de aminoácidos como en las SEC ID 7, 8, 11 ó 12 o la secuencia de los aminoácidos número 1 a 234 de la SEC ID 7 u 8;

b) una proteína que tiene una secuencia de aminoácidos que tiene una similitud de al menos el 90% con la secuencia de las SEC ID 7, 8, 11 ó 12; y

c) una proteína que tiene una secuencia de aminoácidos que tiene una similitud de al menos el 90% con la secuencia de los aminoácidos número 1 a 234 de la SEC ID 7 u 8;

produciendo la expresión del polinucleótido en un sistema que también comprende un receptor nuclear sensible.

3. El uso de acuerdo con la reivindicación 2, en un ensayo para la unión de la proteína con el receptor nuclear sensible.

4. El uso de acuerdo con la reivindicación 2 ó 3, **caracterizado** porque el receptor nuclear sensible es un receptor de progestágenos, un receptor de estrógenos o un receptor de glucocorticoides.

5. Un sistema, que no sea humano, en el que se introduce un polinucleótido y puede producirse su expresión y que además comprende un receptor nuclear sensible, **caracterizado** porque el polinucleótido codifica una proteína selecciona del grupo constituido por:

a) una proteína que tiene una secuencia de aminoácidos como en las SEC ID 7, 8, 11 ó 12 o la secuencia de los aminoácidos número 1 a 234 de la SEC ID 7 u 8;

b) una proteína que tiene una secuencia de aminoácidos que tiene una similitud de al menos el 90% con la secuencia de las SEC ID 7, 8, 11 ó 12; y

c) una proteína que tiene una secuencia de aminoácidos que tiene una similitud de al menos el 90% con la secuencia de los aminoácidos número 1 a 234 en la SEC ID 7 u 8.

6. Un método *in vitro* para determinar el grado de modulación de la actividad transcripcional promovida por un receptor nuclear sensible y un coactivador seleccionado del grupo constituido por:

a) una proteína que tiene una secuencia de aminoácidos como en las SEC ID 7, 8, 11 ó 12 o la secuencia de los aminoácidos número 1 a 234 en la SEC ID 7 u 8;

b) una proteína que tiene una secuencia de aminoácidos que tiene una similitud de al menos el 90% con la secuencia de las SEC ID 7, 8, 11 ó 12; y

c) una proteína que tiene una secuencia de aminoácidos que tiene una similitud de al menos el 90% con la secuencia de los aminoácidos número 1 a 234 de la SEC ID 7 u 8;

en un sistema de acuerdo con la reivindicación 5, que comprende:

a) la adición de un agente al sistema;

b) cuantificar el grado de modulación.

7. El método de acuerdo con la reivindicación 6, **caracterizado** porque el receptor nuclear sensible es un receptor de progestágenos, un receptor de estrógenos o un receptor de glucocorticoides.

8. Un método para la exploración *in vitro* de agentes para influir en la unión entre un receptor nuclear sensible y un coactivador seleccionado del grupo constituido por:

ES 2 305 116 T3

a) una proteína que tiene una secuencia de aminoácidos como en las SEC ID 7, 8, 11 ó 12 o la secuencia de los aminoácidos número 1 a 234 en la SEC ID 7 u 8;

b) una proteína que tiene una secuencia de aminoácidos que tiene una similitud de al menos el 90% con la secuencia en las SEC ID 7, 8, 11 ó 12; y

c) una proteína que tiene una secuencia de aminoácidos que tiene una similitud de al menos el 90% con la secuencia de los aminoácidos número 1 a 234 en la SEC ID 7 u 8;

que comprende comparar la unión entre el receptor nuclear sensible y el coactivador en presencia y en ausencia del agente.

9. El método de acuerdo con la reivindicación 8, **caracterizado** porque el receptor nuclear sensible es un receptor de progestágenos, un receptor de estrógenos o un receptor de glucocorticoides.

10. Un método *in vitro* para el diagnóstico de enfermedades relacionadas con receptores nucleares sensibles que comprende realizar un ensayo para determinar la hibridación cuantitativa con ARN con un polinucleótido marcado o con fragmentos del mismo, **caracterizado** porque el polinucleótido tiene una secuencia que tiene una similitud de al menos el 90% con una o más de las secuencias de la lista que consiste en las SEC ID 1, 2, 3, 4, 5, 6, 9 y 10, y en la que el porcentaje de homología se define, después del alineamiento óptimo de las secuencias usando el algoritmo de Needleman y Wunsch, como el número de restos idénticos entre las dos secuencias dividido por la resta de la longitud de secuencia alineada menos la longitud de todos los huecos.

Figura 1

-> Fragmento de RACE 1
 GGAAGCCCGGGAGTGAGAGAAAGCGGCTCGGGGGCATAGCGGGCCAGTAAGGGCCGCTC
 CTCCTTTGAGAGGTTTGGCGTCTCTTTCCGCCGGTGGCGCTCGGCGCTCACGCAGGGGCG

-> Fragmento de RACE 2
 121 GGTCCCGGTAGCGCCAGGCGGTGTCAGGGCGGGAAGGGGATTGCTGGCGACGGCGGCGGCA
 181 AGGGACAGCAGGAGCAGTGGTGTGTGTCAGCGGGCCGTGGAGACATGGGAGACCCGGGG
 1 M G D P G
 241 TCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACCAACGGAT
 6 S E I I E S V P P A G P E A S E S T T D
 301 GAAATGAAGACGACATTCACTTTGTAGTGAAGGACCATTACGACCTGTTCTTGAATAC
 26 E N E D D I Q F V S E G P L R P V L E Y
 361 ATTGATCTGGTCAGCAGTGTATGAAGAGCTAGCACCTCTTATACTGATGAGAATATT
 46 I D L V S S D D E E P S T S Y T D E N I
 421 AAACGTAAAGACCATATTGATTATCAGAAGGATAAAGTTGCTTAACTCTGGCTCGTCTA
 66 K R K D H I D Y Q K D K V A L T L A R L
 481 GCGCGCCATGTTGAAGTGGAGAAACAGCAGAAAGAAGAGAAGAATAGAGCATTCAAGAGAA
 86 A R H V E V E K Q Q K E E K N R A F R E
 541 AAAATTGATTTTCAGCATGCTCATGGGTACAGAATTTGAATTTATTCGAGGACATTCT
 106 K I D F Q H A H G L Q E L E F I R G H S
 601 GATACAGAAGCAGCAGACTGTGTGTGGACCACTGGCTAAAAATGCCAGGACTCAAAACA
 126 D T E A A R L C V D Q W L K M P G L K T
 661 GGCACATTAATTGTGAACAAAAGTTTCATTCCGAAGAGGAGGCCACACGTGGGTGTCT
 146 G T I N C G T K S S F R R G G H T W V S
 721 GGGAAACCAACTTTATGTCTTATATGCACGTGAACAAGGAGTTTGACAATGGGCACCTT
 166 G K P T L C F I M H C N K E F D M G H L
 781 CTCTTAGACATTTGAAAAGGTTTCGATCACTCTCCATGTGATCCAACATTAACATACAT
 186 L L G H L K R F D H S P C D P T I T L H
 841 GGACCTTTCTTCAGCTCCTTTGCTTGTGTAGTATGTTATAAAAAATTTGTTACTCAACAA
 206 C A P F F S S F A C V V C Y K K F V T Q Q
 901 CAATATAGAGATCACCTTTTTGATAAGGAAGCCACAGATGATGGACATAACAACACACTT
 226 Q Y R D H L F D K E A T D D G H N N N L
 961 CTTCCTCAGATTATTCAGTGTGTTGTCATGTCCAAATGCTTCCTTCTTTTAGCAGAAAG
 246 L P Q I I Q C F A C F N C F L L F S R K
 1021 GAGGAGTGTTCAAAGCATATGTCTGGAAGAATCAATTCATCAGAGTTTCAAACGCGGT
 266 E E C S K E M S G K N H F H Q S F K L G
 1081 GATAACAAAGGAATGCACATCCAATATCTTTCCCATCTTTTGCAAGAAACTTTTGATC
 286 D N K G I A H P I S F P S F A K K L L I
 1141 TCTCTGTCAAAGATGTTCCCTTTCAAGTTAAGTGTGTGGCTGCCACAAGACACTGCGT
 306 S L C K D V P F Q V K C V A C M K T L R
 1201 TCCCACATGGAGCTCACTGCCCATTTCAGAGTTCATTGTGCAAAATGCTGGACCTGTAGCT
 326 S H M E L T A H F R V H C R N A G P V A
 1261 GTAGCTGAGAAGAGCATTACCCAGGTGCCAGAGAAATTCATATTAAGAGGTTATTGTCCA
 346 V A E K S I T Q V A E K F I L R G Y C F
 1321 GATTGCAATCAAGTCTTTGTGGATGAACACAGCAGCCAAAATCATAAGCAGAATTCAGGA
 366 D C M Q V F V D E T S T Q N H K Q M S G
 1381 CACAAAGTCCGAGTCATTAACCTAGTGAAGAATCAGTCTTACTCTATTGCCACAGCAGC
 386 E K V R V I N S V E E S V L L Y C H S S
 1441 GAAGGAAACAAGGATCCTTCTCTGACTTGCAATTTATTGTTGGATCAATCAAAATTTTCA
 406 E G N K D P S S D L H L L L D Q S K F S
 1501 TCACTTAAAGAACCATGTCTATTAAAGAATCTAGCTCACTGGAGTGCAATGGCATTCCA
 426 S L K R T M S I K E S S S L E C I A I P
 1561 AAAAAGAGATGAATTTAAAGATAAAAGCCATGAAGGTGTTGCTTGTGTCCAGAAAGAA
 446 K K K M N L K D K S H E G V A C V Q K E
 1621 AAATCAGTAGTTAAACCTGGTCTGTGAATGCAATCAGCGATTCCCAAGTGAAGATGCA
 466 K S V V K T N F C E C N Q R F P S E D A
 1681 GTAGAAAAGCATGTTTTCTCAGCAACACAATGGGTATAAATGTGTGGTCTGTGGAAG
 486 V E K H V F S A N T M G Y K C V V C G K
 1741 GTATGTGATGATTGAGGGGTCATTGCTTACACATGAGCCGGATTACGGGAGGGGCACAT

Figura 1 continuación

506 V C D D S G V I R L E M S R I H G G A H
 1801 TTAAATAACTTTCTTTTCTGGTGTGCGACATGCAAAAAGGAGTTAACAGGAAGATACT
 526 L N N F L F W C R T C K K E L T R K D T
 1861 ATCATGGCAGATGTGACTGAATTTTCATAATGGACACAGATATTTTATGAGATGGATGAG
 546 I M A M V T E F E N G H R Y F Y E N D E
 1921 GTAGAAGGTGAAACTTTGCCATCATCCTCTACAACATTGGATAATTTGACTGCTAACAG
 566 V E G E T L P S S S T T L D N L T A N K
 1981 CCTTCATCAGCTATTACTGTTATTGATCATTCCTCCCGGCAATAGTTCTCCGAGGGGTA
 586 P S S A I T V I D H S P A N S S P R G K
 2041 TGGCAATGCCGATTGTGAGATATGTTTGATTCCAGGAATATGTAAACAGCAGCTGC
 606 W Q C R I C E D M F D S Q E Y V K Q H C
 2101 ATGTCTTTGGCAAGCCACAAGTTTCATAGATACAGCTGTGCTCACTGCAGAAAGCCTTT
 626 M S L A S H K F H R Y S C A E C R K F F
 2161 CATAAGATAGAACATTGTACCGACATTGCCAAGATGAGCATGACAAATGAGATAAGATT
 646 H K I E T L Y R E C Q D E E D N E I K I
 2221 AAATACTTCTGTGGGCTTTGTGATCTTATCTTTAATGTGGAAGAAGCATTCTGAGTCAT
 666 K Y F C G L C D L I F N V E E A F L S E
 2281 TATGAGGAGCACCACAGCATAGATTATGTTTGTGTGAGAAAAACTGAAACTTCAATT
 686 Y E E E S I D Y V F V S E K T E T S I
 2341 AAACCGAAGATGATTTTCCAGTAATAGAGACCAGTAACCAGTTAATTGTGGTTGCCGT
 706 K T E D D F P V I E T S N Q L T C G C R
 2401 GAGAGTTACATCTGTAAAGTCAACAGAAAAGAGATTATAGCAGATGTCTCCAATCATG
 726 E S Y I C K V N R K E D Y S R C L Q I N
 2461 CTGGATAAAGGAAACTGTGTTTCCGCTGCAGTTTATGTTCCGCAACAGCAGAGATT
 746 L D K G K L W F R C S L C S A T A Q N L
 2521 ACCGACATGAACACTCATATCCATCAAGTGCACAAAGAAAGAGTGATGAGGAGGAGCAG
 766 T D M N T H I R Q V H K E K S D E E E Q
 2581 CAGTATGTAATCAAGTGTGGCACTGCACCAAGCATTTCATGATCCTGAGAGTGCACAG
 786 Q Y V I K C G T C T E A F H D P S R A Q
 2641 CAGCATTTCATAGAAAACATTGCTTCTTACAGAAACCCAGTGTGGCTCATTTTGGATCT
 806 Q H F E R K E C F L Q K P S V A H F G S
 2701 GAAAAATCAACCTGTACAGTTTACTGCTAGTGCCTCACATACAGAGAGAAACTGAA
 826 E K S N L Y K F T A S A S H T E R K L K
 2761 CAGGCAATAAATACTTCAAAAAGTTTAGACATGGAGAAAGGAGTTGAGAATGACCTAAGC
 846 Q A I N Y S K S L D M E K G V E N D L S
 TATCAGAATAT...
 866 Y Q N L...
 variante de corte y empalme: inserción de 144 pares de bases (número de acceso AB011148)
 ..AGAGGAAGAAATTGTTGAGCTTCCAGATTGGGATTACCTGCGAACCATGACTCATATAGT
 ..E E E I V E L P D L D Y L R T N T H I V
 CTTTGTAGATTTTGATAACTGGTCAAACCTTTTTTGGTCATCTACAGGGCATCTAAACCA
 F V D F D N W S N F F G H L P G H L N Q
 AGGAACATTTATTGGGGCTTCA...
 G T F I W G F Q...
 2832 ..AGGAGGAAACACCAATTGGAAGCCTCCGCTCAACTGTAAGATTTATAAC
 870 ..G G N T N W K P P L N C K I Y N
 2881 TACCTGAACAGGATTGGATGCTTCTCCTTCATCCTCGCTGTAGTAAAGAAAAGATGCT
 886 Y L N R I G C F F L H P R C S K R K D A
 2941 GCTGATTTTGCCATATGTATGCTGCTGGCCGTCTAGATGAACAACTACCCAAGCAAT
 906 A D F A I C M H A G R L D E Q L P K Q I
 3001 CCTTTCACCATCCTCTCAGGAGATCAAGGTTTTCTGGAGCTAGAGAATCAATTTAAGAAG
 926 P F T I L S G D Q G F L E L E N Q F K K
 3061 ACTCAGAGGCCAGCTCATATACTAAACCCTCACCCTAGAGGGAGATGATGTGTGCC
 946 T Q R P A H I L N P H H L E G D M M C A
 3121 TTGTTAAATAGCATATCTGATACCACCAAGAAATGTGACAGTGATGATAACATGGGTGCC
 966 L L N S I S D T T K E C D S D D N M G A
 3181 AAAATACTTCAATAGGAGAAGAATTATATCCACAGAAGATGTGGAATTAGAAGAAGCT
 986 K N T S I G E E F I S T E D V E L E A
 3241 ATTAGAAGAAGTCTTGAGGAATGTAATTAAGATATTACCACACAACATCAAGTGGCCT
 1006 I R R S L E E M

Figura 1 continuación

```

3301   TGAAGAGACTGAGATAACGAATTCTTGAGTTTGTCTTCTAAAGGAGACCAGAAATCCACT
3361   ATTACAAATGTATTTGAAAACATGTTTTGCTTTTCATATGTTCAAAATTTGATCTTTGTT
3421   TTGTATTTTTGTGCTAATGTGCAAACATGTACAGAAAGAAATAGAATACATGTTTCATGCAA
3481   ATATAAATTATGTATTTCTAATATAGGTGTAACAGTTTCCCAGTTAACTTTGAATTTATAT
3541   ATTTAGATTTAAAGGATTAAAAAAAGGAAAGCTCTTGACAGTTGTTTCCCAAATAGCAT
3601   TAGTTCTTTAATTTTATTTGTACTGTACAAATGATGCTAGTTTATTTTTTTAGCAGTGAA
3661   AATGAATATAAAAAAGGTGACATAGGTCAAGTTTTCCATAAATCTACTTCTCATGTGGC
3721   ACTATTATATAATACCTTTGAGATCTTTGTTTTATGTTTATTAGAAGAGGGTAATGGAAC
3781   TGTTCATGGAACATTTAAATTTTATCAGCAATGTTCTTTGTGGGTGCCATGCATATGT
3841   AAATTATTGCTTTAATTAGAGCAGCATTGCTAAAATAAAAAATACACCTTAATTTTTTA
3901   AGAATAAATTAATGAACAATTTCCATTCCCTTTACCTGCTTAGACTTTTATGTGACT
3961   TGTATGGTCTCCTGGTTAAAGGGAATGGTGTGAGAAATATTTGCATAAAATTATTTTCAT
4021   AACAGTCCTTTTTTATATATTGGTGTAAACATAATATTCTTCAAGAAATAATGTTGAAAG
4081   CCTACTCAGGAATAATCTTCCTTAACTCTTTAAATTTTTATTTTCATGTGAAGTGTTCAT
4141   GTTTAGTTATTCACTGAATTGTGAGTTTCCTTCATTGGTATTACATATTTAATTCTTAA
4201   ATATTGAGGCCCATTTGTGAGTAATAAAAAATACGACATAAGTAGAAGACTAAGAAGGGC
4261   TTGTCATATTATCTTTGTGTATATGCTTCATGTTATTTAACCAGAAATGCTTTATCTCTA
4321   AAAATTTTATAGTAGTAATTATTTACCAACTTAAAAACATTTCTATAAATATAAAGCTTT
4381   TCTTTATATTTAAGACAAAATATAAAGGCTAGAATTTGGTTCCTTCTCTGTAACACTAAA
4441   TATTTTGTGTGAAATTTGAATTTTTTATTACTATAGTCTTTTCAATCTATAATTGTGT
4501   TTTTAAATTTCTTGATATGTCACTTCTGTTCCCTCCCTGCTTGCATTTTTAAAAACTAGAC
4561   ATTTGTGCTTGAGAAATTTATCAATCATCTTTCCGAAATGACCACATTGTGCTTTTAGTTT
4621   GTATGTTTAAAGTGGTCAAGCAAGGCTCTAGGAGGTAGTCACTGAGCTGGACCTTAAACAC
4681   ATCTGCAGGAGCTCACAAAATGGGAGCAAAAGAGGCATTCCAGAAAGAAGAGTAGTTAAG
4741   CAGATGTGGGGGCAGGGCAAGAGGAGACAGCACCTTGTGCCAGAGATTCTACACAAAGG
4801   TGTGATGGACACTGATGCTCTATTAAAGAGCTTTTGTGGTGTGTGTGGTAGAGAATAATT
4861   AAGGCTTCTGATAAGACAGGAAGGGATATTTTCATCATGACTGTTAAGAAAAGGTAAGTG
4921   GGTGTTTAAATGTTTTAATTGATAATTGTCACTGATCGTATTTCTACTTGTAAATTAAT
4981   AAAATCATCTTAAATTC

```

Figura 2

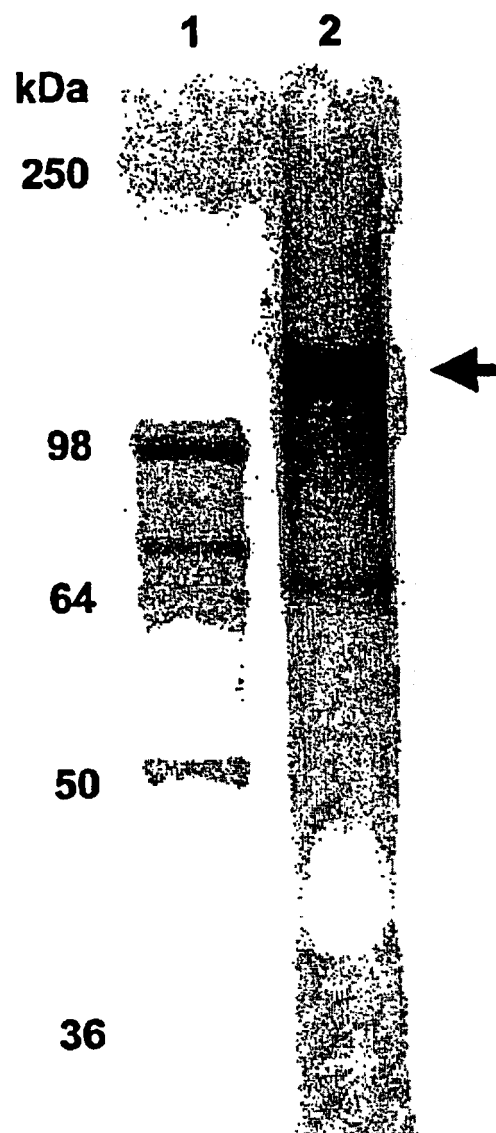


Figura 3

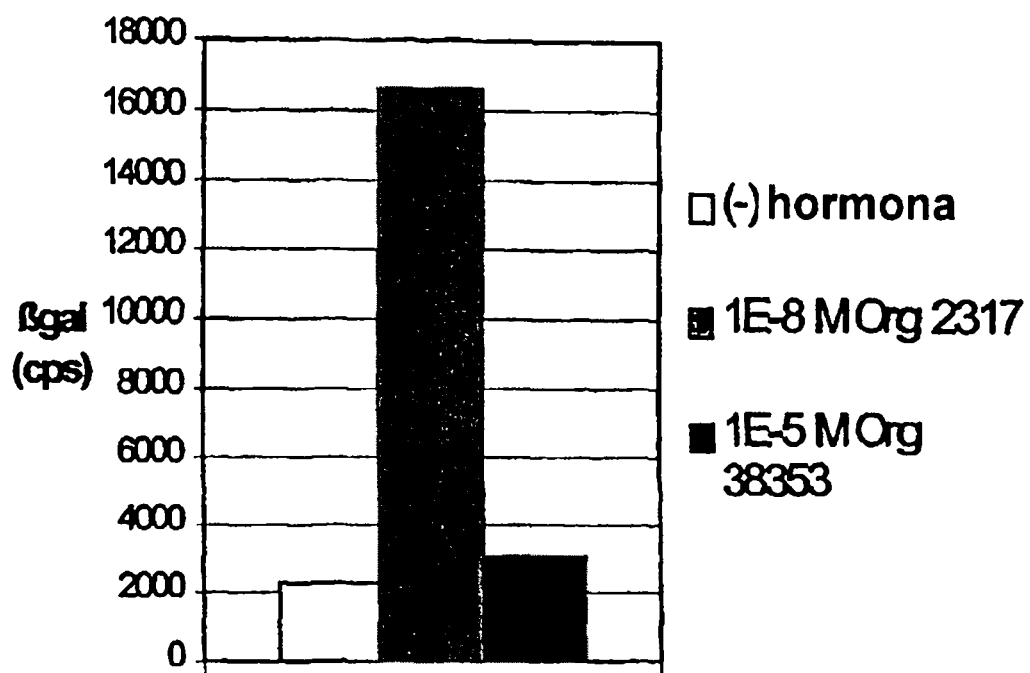


Figura 4

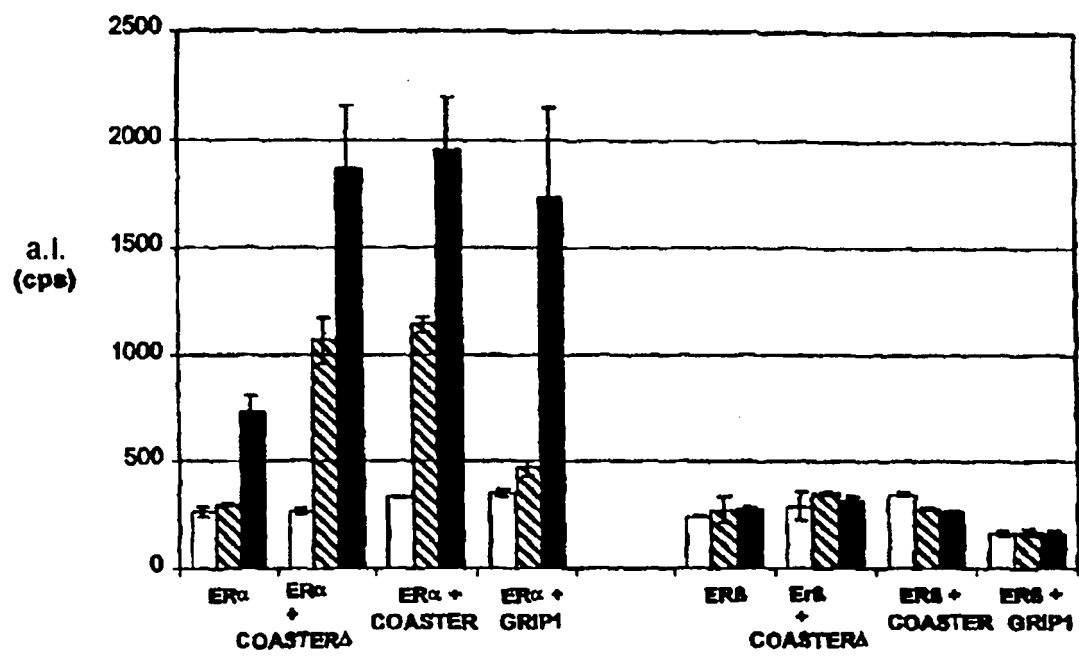


Figura 5

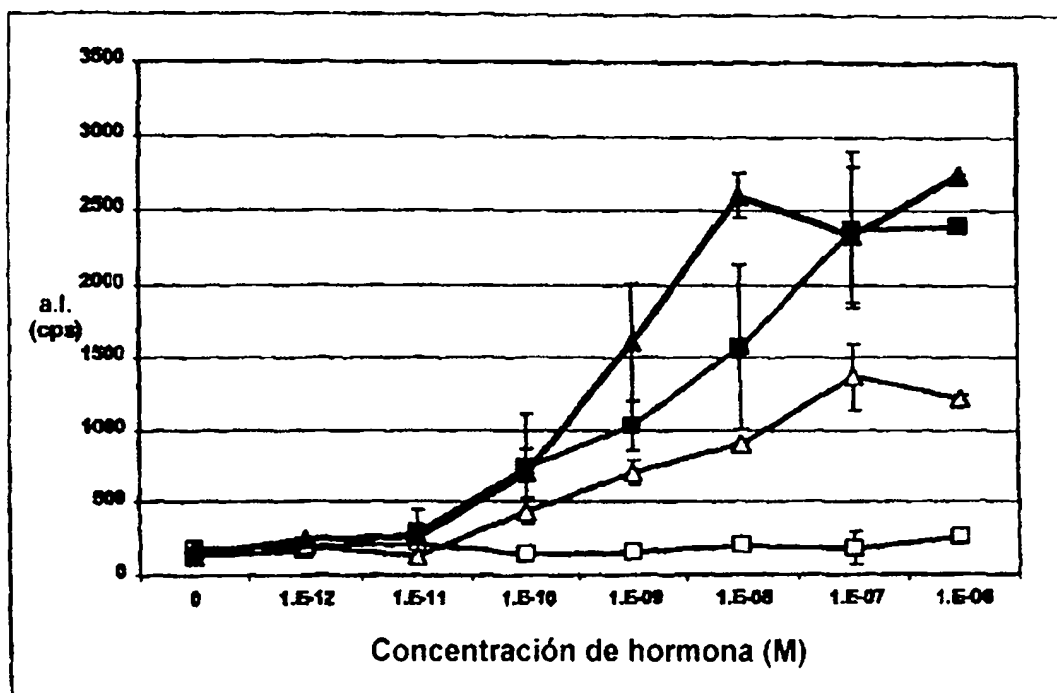


Figura 6

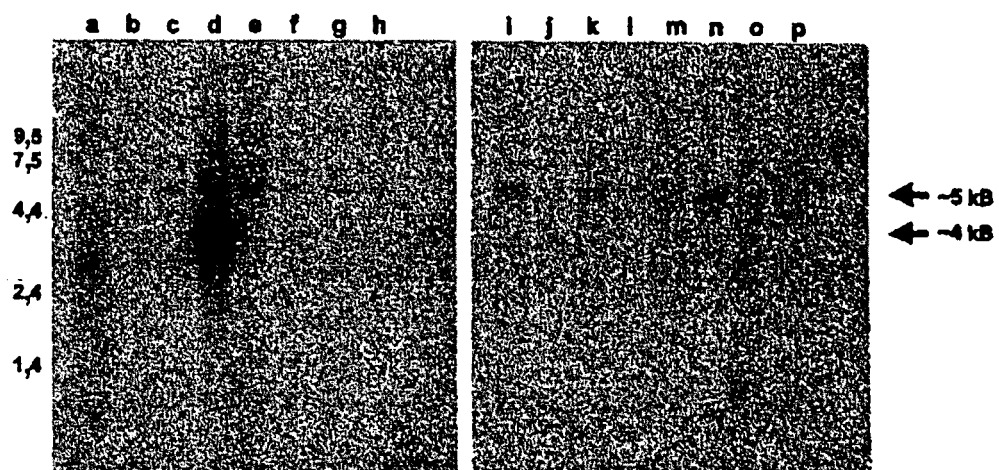


Figura 7

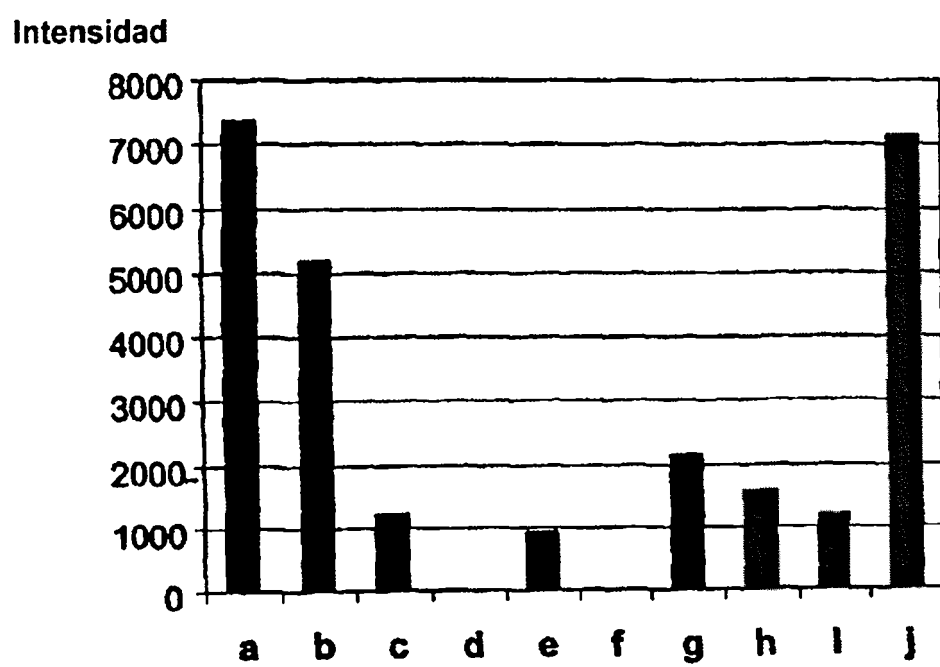


Figura 8

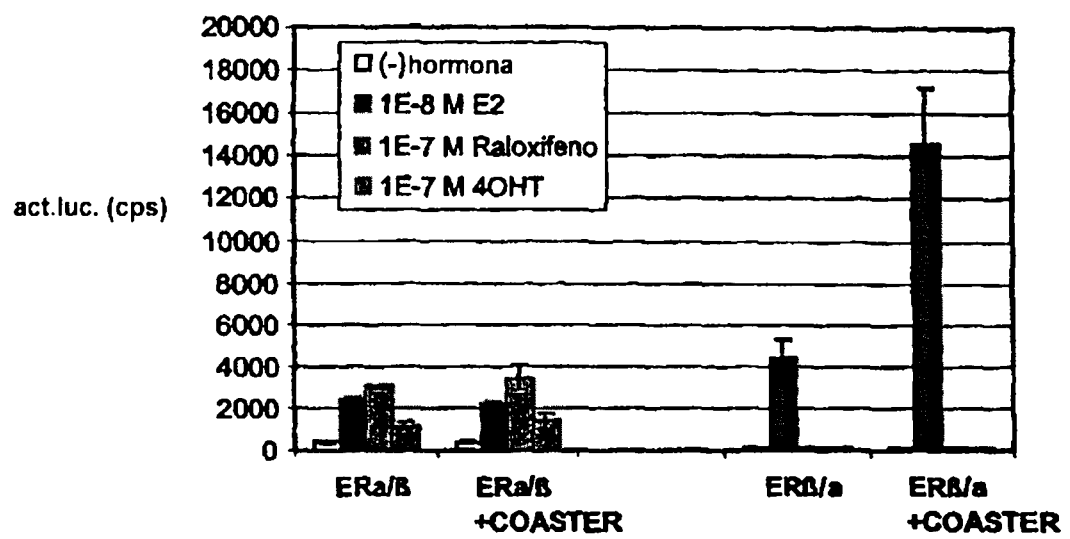


Figura 9

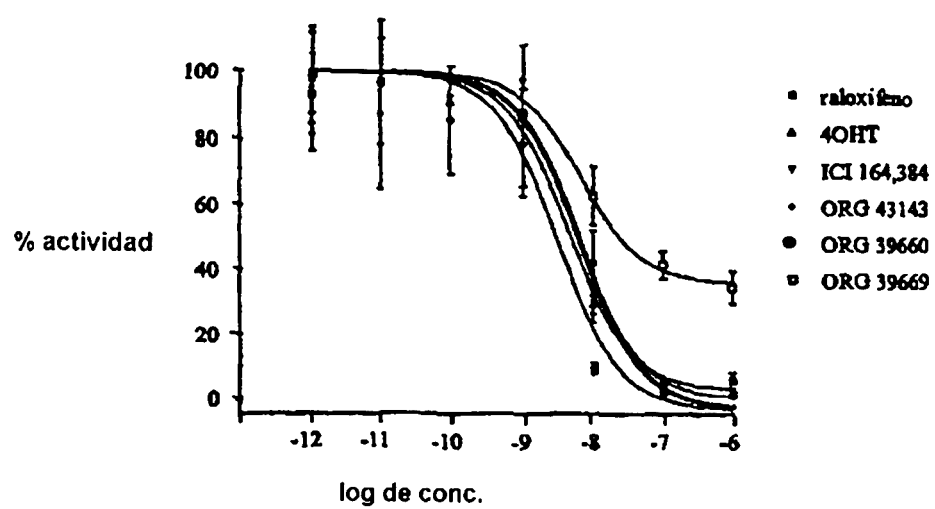


Figura 10

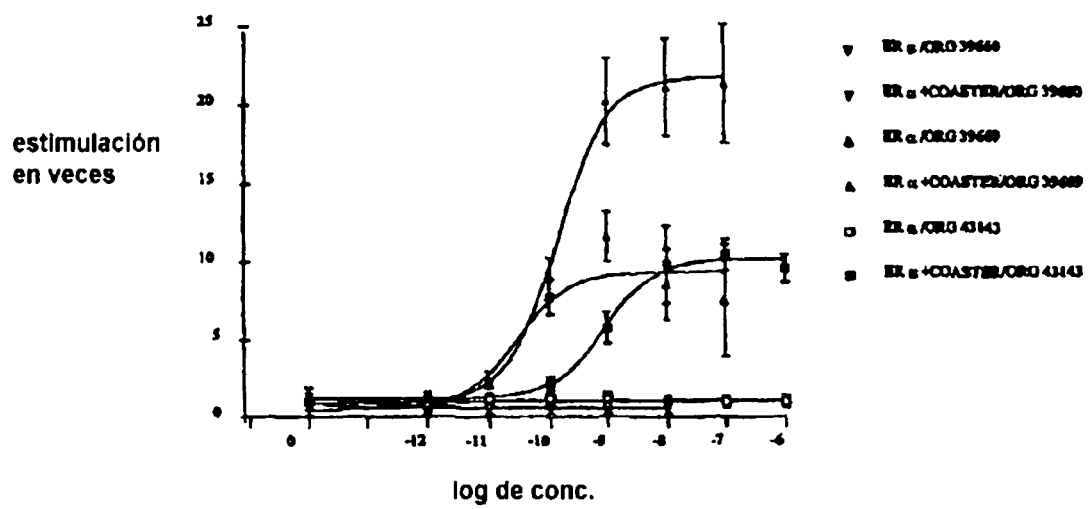
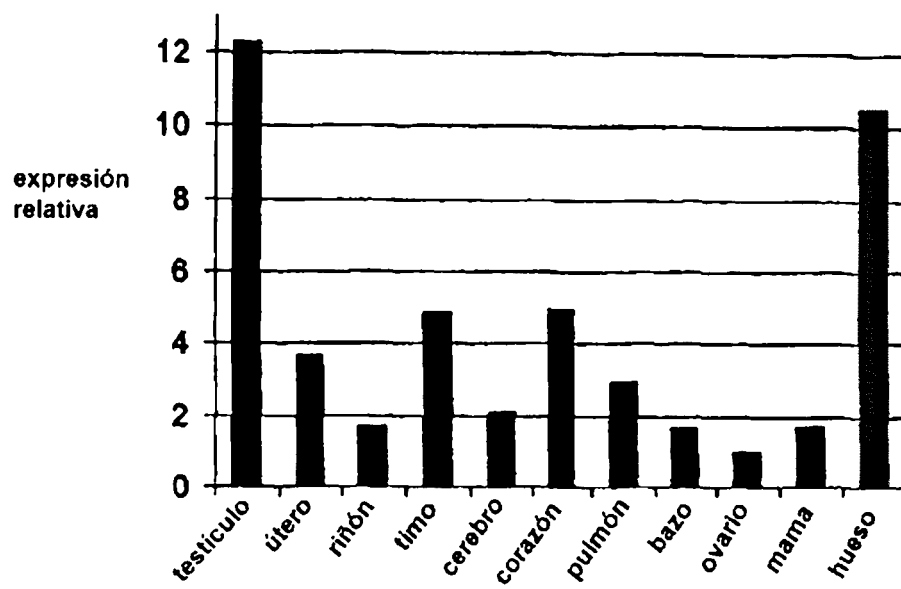


Figura 11



ES 2 305 116 T3

LISTA DE SECUENCIAS

<110> Akzo Nobel N.V.

5 <120> COACTIVACIÓN DE RECEPTORES NUCLEARES

<130> 2000, 569WO

10 <140>

<141>

15 <150> EP202905.6

<151> 21-08-2000

<150> EP1201771.1

20 <151> 14-05-2001

<160> 16

25 <170> PatentIn Ver. 2.1

<210> 1

<211> 4999

30 <212> ADN

<213> *Homo sapiens*

<400> 1

35

```

ggaagcccgaggagtgagagaaagcggctccgggggcatagcgggccagtaagggccgctc 60
ctccttttgaaagggttttgcgtctctttccgccggtggcgtcggcgctcagcagggggcg 120
ggtcccggtagcgcgccaggcggtgcaggggcggaaggggatctgtggcgacggcggcgga 180
agggacagcaggagcagtggtgctgtcagcgcgccgctcgagacatgggagacccgggg 240
tcggagataatagaatctgtccctccagctggccctgaggcatctgagtcacaacgggat 300
gaaaatgaagacgacattcagtttgtcagtgaggaccattacgacctgttcttgaatac 360
attgatctggtcagcagtgatgaagagcctagcacctcttatactgatgagaatatt 420
aaacgtaaagaccatattgatcagaaagataaagttgctttaactctggctcgtcta 480
gcccggcatgttgaagtggagaaacagcagaaagaagagaagaatagagcattcagagaa 540
aaaattgatttcagcatgctcatgggttaagaattggaaattattcgaggacattct 600
gatacagaagcagcaagactgtgtgtggacagtggtgctaaaatgccaggactcaaaaca 660
ggcacaattaattgtggaacaaaaagttcattccgaagaggaggccacacgtgggtgtct 720
gggaaaccaaatttatgtctataatgcactgtaacaaggagtttgacaaagggcacctt 780
ctcttaggacatttgaaaaggttcgatcacctccatgtgatccaacaattacactacat 840
ggacctttcttcagctccttgcttgtgtatgtgttataaaaaatttgtactcaacaa 900
caatatagagatcaccttttgataagggaaccacagatgatggacataaacacaacctt 960
cttcctcagattattcagtttttgcatgtccaaattgcttccttcttttagcagaaag 1020
gaggagtgtcaaagcatatgtctggaaagaaatcatttccatcagagtttcaaactgggt 1080
gataacaaaggattgcacatccaatatcttcccatctttgcaaagaaacttttgatc 1140
tctctgtgcaagatgttccctttcaagtaagtgtgtggcctgccacaaagacactgcgt 1200
tccacatggagctcactgcattttcagagttcattgtcgaaatgctggacctgtagct 1260

```

65

ES 2 305 116 T3

gtagctgaga agagcattac ccaggttgca gagaaattca tattaagagg ttattgtcca 1320
 gattgcaatc aagtccttgt ggatgaaacc agcaccocaa atcataagca gaattcagga 1380
 5 cacaaggtcc gagtcattaa ctccagtggaa gaatcagtc tctctattg ccacagcagc 1440
 gaagggaaca aggatccctc ttctgacttg cttttattgt tggatcaatc aaaattttca 1500
 tcaactaaaa gaaccatgtc tattaagaa tctagctcac tggagtgcac tgccattcca 1560
 aaaaagaaga tgaattttaa agataaaagc catgaagggtg ttgcttgtgt ccagaaagaa 1620
 10 aatcagtag ttaaaacctg gttctgtgaa tgcaatcagc gattcccaag tgaagatgca 1680
 gtagaaaagc atgttttctc agcaaacaca atgggttata aatgtgtggt ctgtggaaag 1740
 gtatgtgatg attcaggggt cattogttta cacatgagcc ggattcacgg aggggcacat 1800
 ttaataact ttcttttctg gtgtcggaca tgcaaaaagg agttaacaag gaaagatact 1860
 15 atcatggcac atgtgactga atttcataat ggacacagat atttttatga gatggatgag 1920
 gtagaagggtg aaactttgcc atcatcctct acaacattgg ataatttgac tgctaacaag 1980
 ccttcatcag ctattactgt tattgatcat tccccggcaa atagttctcc gaggggtaaa 2040
 tggcaatgcc ggatttgtga agatatgttt gattcccgag aatatgtaaa acagcactgc 2100
 20 atgtcttttg caagccacaa gtttcataga tacagctgtg ctactgcag aaagcctttt 2160
 cataagatag aaacattgta ccgacattgc caagatgagc atgacaatga gataaagatt 2220
 aaatacttct gtgggctttg tgatcttate tttaatgtgg aagaagcatt tctgagtcac 2280
 tatgaggagc accacagcat agattatgta tttgtgtcag aaaaaactga aacttcaatt 2340
 25 aaaaacgaag atgattttcc agtaatatag accagtaacc agttaacttg tggttgccgt 2400
 gagagttaca tctgtaaagt caacagaaaa gaagattata gcagatgtct ccaaatcatg 2460
 ctggataaag gaaaactgtg gtttcgctgc agtttatgtt cggcaacagc acagaattta 2520
 30 accgacatga aactcatat ccatacagtg cacaagaaa agagtgatga ggaggagcag 2580
 cagtatgtaa tcaagtgtgg cacctgcacc aaagcatttc atgatcctga gagtgcacag 2640
 cagcatttcc atagaaaaca ttgcttctta cagaaacca gtgtggctca ttttgatct 2700
 gaaaaatcaa acctgtacaa gtttactgct agtgcctcac atacagagag aaaactgaaa 2760
 35 caggcaataa actattcaaa aagtttagac atggagaaa aggttgagaa tgacctaac 2820
 tatcagaata taggaggaaa caccaattgg aagcctccgc tcaactgtaa gatttataac 2880
 tacctgaaca ggattggatg cttcttctct catcctcgtg tagtaaaaag aaaagatgct 2940
 gctgattttg ccataatgtat gcatgctggc cgtctagatg acaactacc caagcaaat 3000
 40 cctttcacca tctctcagg agatcaagg tttctggagc tagagaatca atttaagaag 3060
 actcagaggc cagctcatat actaaacct caccacttag agggagatat gatgtgtgcc 3120
 ttgttaaata gcatatctga taccaccaa gaatgtgaca gtgatgata catgggtgcc 3180
 aaaaatactt caataggaga agaatttata tccacagaag atgtggaatt agaagaagct 3240
 45 attagaagaa gtcttgagga aatgtaatta aagatattac cacacaacat caagtggcct 3300
 tgaagagact gagataacga attcttgagt ttgttttcta aaggagacca gaaatccact 3360
 attacaaatg ttttgaaaa catgtttttg ctttcatatg ttcaaaattt gatctttgtt 3420
 ttgtattttt gtgctaattgt gcaaacatgt acagaagaaa tagaatacat gttcatgcaa 3480
 50 atataaatta tgtattctaa tatagggtga acagtttccc agttaacttt gaatttatat 3540
 atttagattt aaaggattaa aaaaaaggaa agctcttgac agttgtttcc caaatagcat 3600
 tagttcttta attttatttg tactgtacaa atgatgctag tttatttttt tagcagtga 3660
 aatgaatata aaaaagggtga cataggtcaa gttttccata aattctactt ctcatgtggc 3720
 55 actattatat aataoctttg agatctttgt ttatgtttat tagaagaggg taatggaaac 3780
 tgatcatgga aacattttaa attttatcag caatgttctt tgtgggtgcc atgcatatgt 3840
 aaattattgc ttttaattaga gcagcacatt gctaaaataa aaaatacacc ttaattttta 3900
 agaaataatt taaatgaaac aatttccatt ccttttacct gcttagactt ttatgtgact 3960
 60 tgtatggtct cctgggttaa gggaatggtg tcagaatatt tgcataaaat ttttttcat 4020
 aacagtcctt ttttatatat tgggtgaaac ataatttct tcaagaaata atgttgaaag 4080
 octactcagg aataatcttc cttaactctt taaattttta tttcatgtga agtgttttca 4140

65

ES 2 305 116 T3

```

gtttagttat tcaactgaatt gtcagtttcc ttcatttggt attacatatt taattcttaa 4200
atattgaggc cccattgtga gtaataaaaa atacgacata agtagaagac taagaagggc 4260
5 ttgtcatatt atctttgtgt atatgcttca tgttatttaa ccagaaatgt ottatctcta 4320
aaaattttta gtagtaatta ttaccacaaac ttaaaaacat ttctataaat ataaagcttt 4380
tctttatatt taagacaaaa tataaaggct agaatttggt tccttctctg taacactaaa 4440
tatttttagtg tgaaattgaa tttttttatt actatagtct ttccaatcta taatttgtgt 4500
10 ttttaatttc ttgatatgtc acttctgttc ctccctgctt gcatttttta aaaactagac 4560
attgtggctt gagaatttat caatcatctt tccgaaatga ccacattgtg cttttagttt 4620
gtatgtttaa gtggtcaagc aaggctctag gaggtagtca ctgagctgga ccttaaacac 4680
atctgcagga gctcacaaaa tgggagcaaa agaggcattc cagaaagaag agtagttaag 4740
15 cagatgtggg ggcagggcaa gaggagacag caccttggtt ccagagattc tacacaaagg 4800
tgtgatggac actgatgctc tattaagaag cttttgtggt gtgtgtggta gagaataatt 4860
aaggcttctg ataagacagg aaagggatat ttcacatga ctgttaagaa aaggtaactg 4920
20 ggttggttaa tgttttaatt gataattgtc actgatcgtt tttctacttg ttaataatt 4980
aaaaatcatc ttaaaattc 4999

```

<210> 2

<211> 4999

<212> ADN

<213> *Homo sapiens*

<400> 2

```

ggaagcccgg gaggtagaga aagcggctcc gggggcatag cgggccagta agggccgctc 60
35 ctcttttgaa gaggttttgc gtctctttcc gccggtggcg tcggcgctca cgcaggggcg 120
ggtcccggta gcgccaggcg gtgcagggcg ggaaggggat tcgtggcgac ggcgggcgca 180
agggacagca ggagcagtgg tgctgtcagc gcggccgctc gagacatggg agaccggggg 240
40 tcggagataa tagaatctgt ccctccagct ggccctgagg catctgagtc aacaacggat 300
gaaaatgaag acgacattca gtttgtcagt gaaggacat tacgacctgt tcttgaatac 360
attgatctgg tcagcagtga tgatgaagag cctagcacct cttatactga tgagaatatt 420
aaacgtaaaag accatattga ttatcagaag gataaagttg ctttaactct ggctcgtcta 480
45 gcccgccatg ttgaagtgga gaaacagcag aaagaagaga agaataagac attcagagaa 540
aaaattgatt ttcagcatgc tcatgggtta caagaattgg aattttattc aggacattct 600
gatacagaag cagcaagact gtgtgtggac cagtggctaa aaatgccagg actcaaaaca 660
ggcacaatta attgtggaac aaaaagttca ttccgaagag gagggccacac gtgggtgtct 720
50 gggaaaccaaa ctttatgtcc tataatgcac tgtaacaagg agtttgacaa tgggcacctt 780
ctcttaggac atttgaaaag gttcgatcac tctccatgtg atccaacaat tacactacat 840
ggacctttct tcagctcctt tgcttgtgta gtatgttata aaaaatttgt tactcaaca 900
caatatagag atcacctttt tgataaggaa gccacagatg atggacataa caacaacctt 960
55 ctctctcaga ttattcagtg ttttgcagt ccaaattgct tccttctttt tagcagaaag 1020
gaggagtgtt caaagcatat gtctggaaag aatcatttcc atcagagttt caaactgggt 1080
gataacaaag gaattgcaca tccaatatct tcccatctt ttgcaaagaa acttttgatc 1140
tctctgtgca aagatgttcc ctttcaagtt aagtgtgtgg cctgccacaa gacactgcgt 1200
60 tcccacatgg agctcactgc ccatttcaga gttcattgtc gaaatgctgg acctgtagct 1260
gtagctgaga agagcattac ccaggttgca gagaaattca tattaagagg ttattgtcca 1320
gattgcaatc aagtctttgt ggatgaaacc agcaccacaa atcataagca gaattcagga 1380
cacaaagtcc gagtcattaa ctcagtggaa gaatcagctt tactctattg ccacagcagc 1440
65 gaagggaaca aggatccttc ttctgacttg catttattgt tggatcaatc aaaattttca 1500

```

ES 2 305 116 T3

tcacttaaaa gaaccatgtc tattaagaa tctagctcac tggagtgcac tgccattcca 1560
 aaaaagaaga tgaattttaa agataaaagc catgaagggtg ttgcttgtgt ccagaaagaa 1620
 5 aaatcagtag ttaaaacctg gttctgtgaa tgcaatcagc gattcccaag tgaagatgca 1680
 gtagaaaagc atgttttctc agcaaacaca atgggttata aatgtgtggt ctgtggaaag 1740
 gtatgtgatg attcaggggt cattcgttta cacatgagcc ggattcacgg aggggcacat 1800
 ttaaataact ttcttttctg gtgtcggaca tgcaaaaagg agttaacaag gaaagatact 1860
 10 atcatggcac atgtgactga atttcataat ggacacagat atttttatga gatggatgag 1920
 gtagaagggtg aaactttgoc atcatcctct acaacattgg ataatttgac tgctaacaag 1980
 octtcatcag ctattactgt tattgatcat tccccggcaa atagttctcc gaggggtaaa 2040
 tggcaatgcc ggatttgtga agatatgttt gattcccagg aatatgtaaa acagcactgc 2100
 15 atgtcttttg caagccacaa gtttcataga tacagctgtg ctactgcag aaagcctttt 2160
 cataagatag aaacattgta ccgacattgc caagatgagc atgacaatga gataaagatt 2220
 aaatacttct gtgggctttg tgatcttctc tttaatgtgg aagaagcatt tctgagtcac 2280
 tatgaggagc accacagcat agattatgta tttgtgtcag aaaaaactga aacttcaatt 2340
 20 aaaaccgaag atgattttcc agtaatagag accagtaacc agttaacttg tggttgccgt 2400
 gagagttaca tctgtaaaagt caacagaaaa gaagattata gcagatgtct ccaaatcatg 2460
 ctggataaag gaaaactgtg gtttcgctgc agtttatgtt cggcaacagc acagaattta 2520
 accgacatga acactcatat ccatcaagtg cacaagaaa agagtgatga ggaggagcag 2580
 25 cagtatgtaa tcaagtgtgg cacctgcacc aaagcatttc atgatcctga gagtgcacag 2640
 cagcatttcc atagaaaaca ttgcttctta cagaaaacca gtgtggtcct ttttggtctc 2700
 gaaaaatcaa acctgtacaa gtttactgct agtgcctcac atacagagag aaaactgaaa 2760
 caggcaataa actattcaaa aagtttagac atggagaaag gaggttgagaa tgacctaacg 2820
 30 tatcagaata taggaggaaa caccaattgg aagcctccgc tcaactgtaa gatttataac 2880
 tacctgaaca ggattggatg cttcttctct catcctcgct gtagtaaaag aaaagatgct 2940
 gctgattttg ccatatgtat gcatgctggc cgtctagatg aacaactacc caagcaaatt 3000
 35 cctttcacca tctctcagg agatcaagggt tttctggagc tagagaatca atttaagaag 3060
 actcagaggc cagctcatat actaaaccct caccacttag agggagatat gatgtgtgcc 3120
 ttgtttaaata gcatatctga taccaccaa gaatgtgaca gtgatgataa catgggtgcc 3180
 aaaaatactt caataggaga agaatttata tccacagaag atgtggaatt agaagaagct 3240
 40 attagaaaga gtcttgagga aatgtaatta aagatattac cacacaacat caagtggcct 3300
 tgaagagact gagataacga attcttgagt ttgttttcta aaggagacca gaaatccact 3360
 attacaaatg tatttgaaaa catgtttttg ctttcatatg ttcaaaattt gatctttgtt 3420
 ttgtattttt gtgctaattg gcaaacatgt acagaagaaa tagaatacat gttcatgcaa 3480
 45 atataaatta tgtattctaa tatagggtga acagtttccc agttaacttt gaatttataat 3540
 atttagattt aaaggattaa aaaaaaggaa agctcttgac agttgtttcc caaatagcat 3600
 tagttcttta atttfatattg tactgtacaa atgatgctag tttatttttt tagcagtgaa 3660
 aatgaatata aaaaagggtga catagggtcaa gttttccata aattctactt ctcatgtggc 3720
 50 actattatat aatacctttg agatctttgt ttatgtttat tagaagaggg taatggaaac 3780
 tgatcatgga aacattttaa attttatcag caatgttctt tgtgggtgcc atgcatatgt 3840
 aaattattgc tttaattaga gcagcacatt gctaaaataa aaaatacacc ttaattttta 3900
 agaaataatt taaatgaaac aatttccatt ccttttacct gcttagactt ttatgtgact 3960
 55 tgtatggctc cctgggttaa gggaatggtg tcagaatatt tgcataaaat tatttttcat 4020
 aacagtcctt ttttatatat tgggtgtaa acataatattc tcaagaaata atgttgaaag 4080
 octactcagg aataatcttc ctttaactctt taaattttta tttcatgtga agtgttttca 4140
 gtttagttat tcaactgaatt gtcagtttcc ttcatttgggt attacatatt taattcttaa 4200
 60 atattgaggc ccatttgtga gtaataaaaa atacgacata agtagaagac taagaagggc 4260
 ttgtcatatt atcttttgtt atatgcttca tgttatttaa ccagaaatgt cttatctcta 4320
 aaaattttta gtagtaatta tttaccaaac ttaaaaacat ttctataaat ataaagcttt 4380

ES 2 305 116 T3

tctttatatt taagacaaaa tataaaggct agaatttggc tccttctctg taacactaaa 4440
 tatttttagtg tgaaattgaa tttttttatt actatagtct tttcaatcta taatttgtgt 4500
 5 ttttaatttc ttgatatgtc acttctgttc ctccctgctt gcatttttta aaaactagac 4560
 attgtggctt gagaatttat caatcatctt tccgaaatga ccacattgtg cttttagttt 4620
 gtatgtttta gtggtcaagc aaggctctag gaggtagtca ctgagctgga ccttaaacac 4680
 atctgcagga gctcacaaaa tgggagcaaa agaggcattc cagaaagaag agtagttaag 4740
 10 cagatgtggg ggcagggcaa gaggagacag caccttggtg ccagagattc tacacaaagg 4800
 tgtgatggac actgatgtct tattaagaag cttttgtggt gtgtgtggta gagaataatt 4860
 aaggcttctg ataagacagg aaagggatat tcatcatga ctgttaagaa aaggtaactg 4920
 gggtgtttta tgttttaatt gataattgtc actgatcgta tttctacttg ttaaataatt 4980
 15 aaaaatcatc ttaaaattc 4999

<210> 3

20 <211> 5144

<212> ADN

<213> *Homo sapiens*

25 <400> 3

ggaagcccg gagtgagaga aagcggctcc gggggcatag cgggccagta agggccgctc 60
 30 ctcttttgaa gaggttttgc gtctctttcc gccgggtggc tcggcgctca cgcaggggcg 120
 ggtcccggt ggcgcaggcg gtgcagggcg ggaaggggat tcgtggcgac ggcggcggca 180
 agggacagca ggagcagtg tgctgtcagc gcggccgctc gagacatggg agaccgggg 240
 tcggagataa tagaatctgt cctccagct ggccctgagg catctgagtc aacaacggat 300
 35 gaaaatgaag acgacattca gtttgtcagt gaaggacat tacgacctgt tcttgaatac 360
 attgatctgg tcagcagtga tgatgaagag cctagcacct cttatactga tgagaatatt 420
 aaacgtaaag accatattga ttatcagaag gataaagttg ctttaactct ggctcgtcta 480
 gcccgccatg ttgaagtga gaaacagcag aaagaagaga agaataagac attcagagaa 540
 40 aaaattgatt ttcagcatgc tcatgggtta caagaattgg aatttattcg aggacattct 600
 gatacagaag cagcaagact gtgtgtggac cagtggctaa aaatgccagg actcaaaaca 660
 ggcacaatta attgtggaac aaaaagttca ttccgaagag gaggccacac gtgggtgtct 720
 gggaaaccaa ttttatgtcc tataatgcac tgtaacaagg agtttgacaa tgggcacctt 780
 45 ctcttaggac atttgaaaag gttcgatcac tctccatgtg atccaacaat tacactacat 840
 ggacctttct tcagctcctt tgcttgtgta gtatgttata aaaaatttgt tactcaacaa 900
 caatatagag atcacctttt tgataaggaa gccacagatg atggacataa caacaacctt 960
 50 ctctctcaga ttattcagtg ttttgcattg ccaaattgct tccttctttt tagcagaaag 1020
 gaggagtgtt caaagcatat gtctggaaag aatcatttcc atcagagttt caaactgggt 1080
 gataacaaag gaattgcaca tccaatatct tcccatctt ttgcaaagaa acttttgatc 1140
 tctctgtgca aagatgttcc ctttcaagtt aagtgtgtgg cctgccacaa gacactgctg 1200
 55 tcccatatgg agctcactgc ccatttcaga gttcattgtc gaaatgctgg acctgtagct 1260
 gtagctgaga agagcattac ccagggtgca gagaaattca tattaagagg ttattgtcca 1320
 gattgcaatc aagtctttgt ggatgaaacc agcaccocaa atcataagca gaattcagga 1380
 cacaaagtcc gagtcatata ctcagtggaa gaatcagctt tactctattg ccacagcagc 1440
 60 gaagggaaca aggatccttc ttctgacttg cttttattgt tggatcaatc aaaattttca 1500
 tcacttaaaa gaacctgtc tattaagaa tctagctcac tggagtgcac tgccattcca 1560
 aaaaagaaga tgaatttaaa agataaaagc catgaaggtg ttgcttgtgt ccagaaagaa 1620
 aaatcagtag ttaaaacctg gttctgtgaa tgcaatcagc gattcccaag tgaagatgca 1680
 65 gtagaaaagc atgttttctc agcaaacaca atgggttata aatgtgtggt ctgtggaaag 1740

ES 2 305 116 T3

gtatgtgatg attcaggggt cattcggttta cacatgagcc ggattcacgg aggggcacat 1800
 ttaaataaact ttcttttctg gtgtcggaca tgcaaaaagg agttaacaag gaaagatact 1860
 5 atcatggcac atgtgactga atttcataat ggacacagat atttttatga gatggatgag 1920
 gtagaagggtg aaactttgcc atcatcctct acaacatttg ataatttgac tgctaacaag 1980
 ccttcacacag ctattactgt tattgatcat tccccggcaa atagttctcc gaggggtaaa 2040
 tggcaatgcc ggatttgtga agatatgttt gattcccgag aatatgtaaa acagcactgc 2100
 10 atgtcttttg caagccacaa gtttcataga tacagctgtg ctactgcag aaagcctttt 2160
 cataagatag aaacatttga ccgacattgc caagatgagc atgacaatga gataaagatt 2220
 aaatacttct gtgggctttg tgatcttctc tttaatgtgg aagaagcatt tctgagtcac 2280
 tatgaggagc accacagcat agattatgta tttgtgtcag aaaaaactga aacttcaatt 2340
 15 aaaaaccgaag atgattttcc agtaatagag accagtaacc agttaacttg tgggtgccgt 2400
 gagagttaca tctgtaaaagt caacagaaaa gaagattata gcagatgtct ccaaatcatg 2460
 ctggataaag gaaaactgtg gtttcgctgc agtttatgtt cggcaacagc acagaattta 2520
 accgacatga acactcatat ccatcaagtg cacaagaaa agagtgatga ggaggagcag 2580
 20 cagtatgtaa tcaagtgtgg cacctgcacc aaagcatttc atgatcctga gagtgcacag 2640
 cagcatttcc atagaaaaca ttgcttctta cagaaaccca gtgtggctca ttttggatct 2700
 gaaaaatcaa acctgtacaa gtttactgct agtgcctcac atacagagag aaaactgaaa 2760
 caggcaataa actattcaaa aagtttagac atggagaaaag gagttgagaa tgacctaacg 2820
 25 tatcagaata tagaggaaga aattgttgag ctccagatt tggattacct gcgaacctag 2880
 actcatatag tctttgtaga ttttgataac tgggtcaaact ttttgggtca tctaccaggg 2940
 catctaaacc aaggaaacatt tatttggggc tttcaaagga ggaaacacca attggaagcc 3000
 30 tccgctcaac tgtaagattt ataactacct gaacaggatt ggatgcttct tcttcatcc 3060
 tgcgtgtagt aaaagaaaag atgctgctga ttttgccata tgtatgcag ctggcctct 3120
 agatgaacaa ctacccaagc aaattccttt caccatcctc tcaggagatc aaggttttct 3180
 ggagctagag aatcaattta agaagactca gaggccagct catatactaa accctacca 3240
 35 cttagagggg gatatgatgt gtgccttggt aaatagcata tctgatacca ccaaagaatg 3300
 tgacagtgat gataacatgg gtgccaaaaa tacttcaata ggagaagaat ttatatccac 3360
 agaagatgtg gaattagaag aagctattag aagaagtctt gaggaatgt aattaaagat 3420
 attaccacac aacatcaagt ggccttgaag agactgagat aacgaattct tgagtttgtt 3480
 40 ttctaaagga gaccagaaat ccactattac aaatgtattt gaaaacatgt ttttgctttc 3540
 atatgttcaa aatttgatct ttgttttgta tttttgtgct aatgtgcaaa catgtacaga 3600
 agaaatagaa tacatgttca tgcaaatata aattatgtat tctaataatag gtgtaacagt 3660
 ttcccagtta actttgaatt tatatattta gatttaaagg attaaaaaa aggaaagctc 3720
 45 ttgacagttg ttcccaaatt agcattagtt cttaaatttt atttgactg tacaaatgat 3780
 gctagtttat ttttttagca gtgaaaatga atataaaaaa ggtgacatag gtcaagtttt 3840
 ccataaattc tacttctcat gtggcactat tatataatac ctttgagatc tttgtttatg 3900
 tttattagaa gagggtaatg gaaactgatc atggaaacat ttaaaatttt atcagcaatg 3960
 50 ttctttgtgg gtgccatgca tatgtaaatt attgctttta ttagagcagc acattgctaa 4020
 aataaaaaat acaccttaat ttttaagaaa taatttaaatt gaaacaattt ccattccttt 4080
 tactgctta gacttttatg tgacttgtat ggtctcctgg ttaaagggaa tgggtgtcaga 4140
 atatttgcac aaaattattt ttcataacag tcctttttta tatattggtg taaacataat 4200
 55 attcttcaag aaataatgtt gaaagcctac tcaggaataa tcttccttaa ctctttaaat 4260
 ttttatttca tgtgaagtgt tttcagttta gttattcact gaattgtcag tttccttcat 4320
 ttggtattac atatttaatt cttaaattat gaggcccat tgtgagtaat aaaaaatacg 4380
 60 acataagtag aagactaaga agggcttgct atattatctt tgtgtatatg cttcatgtta 4440
 ttttaaccaga aatgtcttat ctctaaaaat ttttagtagt aattatttac caaacttaaa 4500
 aacatttcta taaatataaa gcttttcttt atatttaaga caaaatataa aggctagaat 4560
 ttggttcctt ctctgtaaca ctaaatattt tagtgtgaaa ttgaattttt ttattactat 4620

ES 2 305 116 T3

```

    agtcttttca atctataatt tgtgttttta atttcttgat atgtcacttc tgttcctccc 4680
    tgettgcatt ttttaaaaac tagacattgt ggcttgagaa tttatcaatc atctttccga 4740
5   aatgaccaca ttgtgctttt agtttgtatg ttttaagtggc caagcaaggc tctaggaggt 4800
    agtcaactgag ctggacctta aacacatctg caggagctca caaaatggga gcaaaagagg 4860
    cattccagaa agaagagtag ttaagcagat gtggggggcag ggcaagagga gacagcacct 4920
    tgttgccaga gattctacac aaaggtgtga tggacactga tgctctatta agaagctttt 4980
10  gtggtgtgtg tggtagagaa taattaaggc ttctgataag acaggaaagg gatatttcat 5040
    catgactgtt aagaaaaggc aactgggttg tttaatgttt taattgataa ttgtcactga 5100
    tcgtatttct acttggttaa taattaataa tcactttaaa attc 5144

```

```

15  <210> 4
    <211> 5144
    <212> ADN
20  <213> Homo sapiens

    <400> 4

```

```

25  ggaagcccgg gagtgagaga aagcggctcc gggggcatag cggggccagta agggccgctc 60
    ctcttttgaa gaggttttgc gtctctttcc gccggtggcg tcggcgctca cgcaggggcg 120
    ggtcccggta gcgccaggcg gtgcagggcg ggaaggggat tcgtggcgac ggcgggcgca 180
30  agggacagca ggagcagtgg tgctgtcagc gcggccgtcg gagacatggg agaccggggg 240
    tcggagataa tagaatctgt ccctccagct ggccctgagg catctgagtc aacaacggat 300
    gaaaatgaag acgacattca gtttgtcagt gaaggaccat taogacctgt tcttgàatac 360
    attgatctgg tcagcagtga tgatgaagag cctagcacct cttatactga tgagaatatt 420
35  aaacgtaaag accatattga ttatcagaag gataaagttg ctttaactct ggctcgtcta 480
    gcccgccatg ttgaagtga gaaacagcag aaagaagaga agaataagagc attcagagaa 540
    aaaattgatt ttcagcatgc tcatgggtta caagaattgg aatttattcg aggacattct 600
    gatacagaag cagcaagact gtgtgtggac cagtggctaa aatgccagg actcaaaaca 660
40  ggcacaatta attgtggaac aaaaagttca ttccgaagag gaggccacac gtgggtgtct 720
    gggaaaccaa ctttatgtcc tataatgcac tgtaacaagg agtttgacaa tgggcacctt 780
    ctcttaggac atttgaaaag gttcgatcac tctccatgtg atccaacaat tacactacat 840
    ggacctttct tcagctcctt tgcttgtgta gtatgttata aaaaatttgt tactcaacaa 900
45  caatatagag atcacctttt tgataaggaa gccacagatg atggacataa caacaacctt 960
    ctccctcaga ttattcagtg ttttgcagt ccaaattgct tccttctttt tagcagaaag 1020
    gaggagtgtt caaagcatat gtctggaaag aatcatttcc atcagagttt caaactgggt 1080
50  gataacaaag gaattgcaca tccaatatct tcccacatt ttgcaaagaa acttttgatc 1140
    tctctgtgca aagatgttcc ctttcaagtt aagtgtgtgg cctgccacaa gacactgcgt 1200
    tcccacatgg agctcactgc ccatttcaga gttcattgtc gaaatgctgg acctgtagct 1260
    gtagctgaga agagcattac ccagggttga gagaaattca tattaagagg ttattgtcca 1320
55  gattgcaatc aagtctttgt ggatgaaacc agcacccaaa atcataagca gaattcagga 1380
    cacaaagtc gagtcattaa ctcagtggaa gaatcagctt tactctattg ccacagcagc 1440
    gaagggaaca aggatccttc ttctgacttg catttattgt tggatcaatc aaaattttca 1500
    tcacttaaaa gaaccatgtc tattaagaaa tctagctcac tggagtgcac tgccattcca 1560
60  aaaaagaaga tgaatttaaa agataaaaagc catgaagggtg ttgcttgtgt ccagaaagaa 1620
    aaatcagtag ttaaaacctg gttctgtgaa tgcaatcagc gattcccaag tgaagatgca 1680
    gtagaaaagc atgttttctc agcaaacaca atgggttata aatgtgtggt ctgtggaaag 1740
    gtatgtgatg attcaggggt cattcggtta cacatgagcc ggattcacgg aggggcacat 1800
65  ttaataaact ttcttttctg gtgtcggaca tgcaaaaagg agttaacaag gaaagatact 1860

```

ES 2 305 116 T3

atcatggcac atgtgactga atttcataat ggacacagat atttttatga gatggatgag 1920
 gtagaagggtg aaactttgcc atcatcctct acaacattgg ataatttgac tgctaacaag 1980
 5 ccttcatcag ctattactgt tattgatcat tccccggcaa atagttctcc gaggggtaaa 2040
 tggcaatgcc ggattttgtga agatatgttt gattcccagg aatatgtaaa acagcactgc 2100
 atgtcttttg caagccacaa gtttcataga tacagctgtg ctactgcag aaagcctttt 2160
 cataagatag aaacattgta ccgacattgc caagatgagc atgacaatga gataaagatt 2220
 10 aaatacttct gtgggctttg tgatcttctc tttaatgtgg aagaagcatt tctgagtcac 2280
 tatgaggagc accacagcat agattatgta tttgtgtcag aaaaaactga aacttcaatt 2340
 aaaaccgaag atgattttcc agtaatatag accagtaacc agttaacttg tggttgccgt 2400
 15 gagagttaca tctgtaaagt caacagaaaa gaagattata gcagatgtct ccaaatcatg 2460
 ctggataaag gaaaactgtg gtttcgctgc agtttatgtt cggcaacagc acagaattta 2520
 accgacatga aactcatat ccatcaagtg cacaaagaaa agagtgatga ggaggagcag 2580
 cagtatgtaa tcaagtgtgg cacctgcacc aaagcatttc atgacctga ggtgcacag 2640
 20 cagcatttcc atagaaaaa ttgcttctta cagaaacca gtgtggctca ttttggatct 2700
 gaaaaatcaa acctgtacaa gtttactgct agtgcctcac atacagagag aaaactgaaa 2760
 caggcaataa actattcaaa aagtttagac atggagaaa gagttgagaa tgacctaac 2820
 tatcagaata tagaggaaga aattgttgag ctccagatt tggattacct ggaacctag 2880
 25 actcatatag tctttgtaga ttttgataac tggcctaaact tttttgtca tctaccagg 2940
 catctaaacc aaggaacatt tatttggggc tttcaaagga ggaaacacca attggaagcc 3000
 tccgctcaac tgtaagattt ataactacct gaacaggatt ggatgctct tcttcatcc 3060
 tctgttagt aaaagaaaag atgctgctga ttttgccata tgtatgcag ctggcctct 3120
 30 agatgaacaa ctacccaagc aaattccttt caccatctc tcaggagatc aaggttttct 3180
 ggagctagag aatcaattta agaagactca gaggccagct catatactaa acctcacca 3240
 cttagagggg gatatgatgt gtgccttgtt aaatagcata tctgatacca ccaagaatg 3300
 tgacagtgat gataacatgg gtgccaaaa tacttcaata ggagaagaat ttatatccac 3360
 35 agaagatgtg gaattagaag aagctattag aagaagtctt gaggaatgt aattaaagat 3420
 attaccacac aacatcaagt ggccttgaag agactgagat aacgaattct tgagtttgtt 3480
 ttctaaagga gaccagaaat ccactattac aaatgtattt gaaaacatgt ttttgcttcc 3540
 atatgttcaa aatttgatct ttgttttcta tttttgtgct aatgtgcaaa catgtacaga 3600
 40 agaaatagaa tacatgttca tgcaaatata aattatgtat tctaatatag gtgtaacagt 3660
 ttcccagtta actttgaatt tatatattta gatttaaagg attaaaaaaa aggaaagctc 3720
 ttgacagttg tttcccaaat agcattagtt ctttaatttt atttgtactg tacaatgat 3780
 gctagtttat ttttttagca gtgaaaatga atataaaaaa ggtgacatag gtcaagtttt 3840
 45 ccataaattc tacttctcat gtggcactat tatataatac ctttgagatc tttgtttatg 3900
 tttattagaa gagggtaatg gaaactgatc atggaaacat ttaaaatttt atcagcaatg 3960
 ttctttgtgg gtgccatgca tatgtaaatt attgctttta ttagagcagc acattgctaa 4020
 aataaaaaat acaccttaat ttttaagaaa taatttaaat gaaacaattt ccattccttt 4080
 50 tacctgctta gacttttatg tgacttgtat ggtctcctgg ttaaagggaa tgggtgtcaga 4140
 atatttgcac aaaattattt ttcataacag tcctttttta tatattggtg taaacataat 4200
 attcttcaag aaataatgtt gaaagcctac tcaggaataa tcttccttaa ctotttaaatt 4260
 55 ttttatttca tgtgaagtgt tttcagttta gttattcact gaattgtcag tttccttcat 4320
 ttggtattac atatttaatt cttaaataat gagggcccat tgtgagtaat aaaaaatacg 4380
 acataagtag aagactaaga agggcttgtc atattatctt tgtgtatatg ctcatgtta 4440
 ttttaaccaga aatgtcttat ctctaaaaat ttttagtagt aattatttac caaacttaa 4500
 60 aacatttcta taaatataaa gcttttcttt atatttaaga caaaatataa aggctagaat 4560
 ttggttctct ctctgtaaca ctaaatattt tagtgtgaaa ttgaattttt ttattactat 4620
 agtcttttca atctataatt tgtgttttta atttcttgat atgtcacttc tgttctccc 4680
 65 tgcttgcatt ttttaaaaac tagacattgt ggcttgagaa tttatcaatc atctttccga 4740

ES 2 305 116 T3

```

aatgaccaca ttgtgctttt agtttgtatg tttaagtggc caagcaaggc tctaggaggt 4800
agtcactgag ctggacctta aacacatctg caggagctca caaatggga gcaaaagagg 4860
5 cattccagaa agaagagtag ttaagcagat gtgggggcag ggcaagagga gacagcacct 4920
tggtgccaga gattctacac aaaggtgtga tggacactga tgctctatta agaagctttt 4980
gtggtgtgtg tggtagagaa taattaaggc ttctgataag acaggaaagg gatatttcat 5040
catgactgtt aagaaaaggc aactgggttg tttaatgttt taattgataa ttgtcactga 5100
10 tctgatttct acttggttaa taattaaaaa tcctcttaaa attc 5144

```

<210> 5

15 <211> 3042

<212> ADN

<213> *Homo sapiens*

20 <400> 5

```

atgggagacc cggggtcgga gataatagaa tctgtccctc cagctggccc tgaggcatct 60
25 gagtcaacaa cggatgaaaa tgaagacgac attcagtttg tcagtgaagg accattacga 120
cctgttcttg aatacattga tctggtcagc agtcatgatg aagagcctag cacctcttat 180
actgatgaga atattaaacg taaagaccat attgattatc agaaggataa agttgcttta 240
actctggctc gtctagcccg ccatgttgaa gtggagaaac agcagaaaga agagaagaat 300
30 agagcattca gagaaaaaat tgattttcag catgctcatg gggtacaaga attggaattt 360
attcgaggac attctgatac agaagcagca agactgtgtg tggaccagtg gctaaaaatg 420
ccaggactca aaacaggcac aattaattgt ggaacaaaaa gttcattccg aagaggaggc 480
cacacgtggg tgtctgggaa accaatttta tgcctataa tgcactgtaa caaggagttt 540
35 gacaatgggc accttctctt aggacatttg aaaaggttcg atcactctcc atgtgatcca 600
acaattacac tacatggacc tttcttcagc tcctttgctt gtgtagtatg ttataaaaaa 660
tttgttactc aacaacaata tagagatcac ctttttgata aggaagccac agatgatgga 720
40 cataacaaca accttcttcc tcagattatt cagtgttttg catgtccaaa ttgcttctct 780
cttttttagca gaaaggagga gtgttcaaag catatgtctg gaaagaatca ttccatcag 840
agtttcaaac tgggtgataa caaaggaatt gcacatccaa tatctttccc atcttttgca 900
aagaaacttt tgatctctct gtgcaaagat gttcccttcc aagttaagtg tgtggcctgc 960
45 cacaagacac tgcgttccca catggagctc actgccattc tcagagttca ttgtcgaaat 1020
gctggacctg tagctgtagc tgagaagagc attaccaggg ttgcagagaa attcatatta 1080
agaggttatt gtccagattg caatcaagtc tttgtggatg aaaccagcac ccaaatcat 1140
aagcagaatt caggacacaa agtccgagtc attaaactcag tggagaatc agtcttactc 1200
50 tattgccaca gcagogaagg gaacaaggat ccttcttctg acttgcatth attgttggtat 1260
caatcaaaat tttcatcact taaaagaacc atgtctatta aagaatctag ctactggag 1320
tgcatthcca ttccaaaaaa gaagatgaat ttaaaagata aaagccatga aggtgttgct 1380
tgtgtccaga aagaaaaatc agtagttaa acctggttct gtgaatgcaa tcagcgattc 1440
55 ccaagtgaag atgcagtaga aaagcatgtt ttctcagcaa acacaatggg ttataaatgt 1500
gtggtctgtg gaaaggtagt tgatgattca ggggtcattc gtttacacat gagccggatt 1560
cacggagggg cacattttaa taactttctt ttctggtgtc ggacatgcaa aaaggagtta 1620
acaaggaaag atactatcat ggcacatgtg actgaatttc ataatggaca cagatatttt 1680
60 tatgagatgg atgaggtaga aggtgaaact ttgcatcat cctctacaac attggataat 1740
ttgactgcta acaagccttc atcagctatt actgttattg atcattcccc ggcaaatagt 1800
tctccgaggg gtaaatggca atgocggatt tgtgaagata tgtttgattc ccaggaatat 1860
gtaaaacagc actgcatgtc tttggcaagc cacaagtttc atagatacag ctgtgctcac 1920
65 tgcagaaagc cttttcataa gatagaaaca ttgtaccgac attgccaaga tgagcatgac 1980

```

ES 2 305 116 T3

aatgagataa agattaaata cttctgtggg ctttgtgatc ttatctttaa tgtggaagaa 2040
 gcatttctga gtcattatga ggagcaccac agcatagatt atgtatttgt gtcagaaaaa 2100
 5 actgaaactt caattaaaac cgaagatgat tttccagtaa tagagaccag taaccagtta 2160
 acttgtgggt gccgtgagag ttacatctgt aaagtcaaca gaaaagaaga ttatagcaga 2220
 tgtctccaaa tcatgctgga taaaggaaaa ctgtggttcc gctgcagttt atgttcggca 2280
 acagcacaga atttaaccga catgaacact catatccatc aagtgcacaa agaaaagagt 2340
 10 gatgaggagg agcagcagta tgtaatcaag tgtggcacct gcaccaaagc atttcatgat 2400
 cctgagagtg cacagcagca tttccataga aaacattgct tcttacagaa acccagtgtg 2460
 gctcattttg gatctgaaaa atcaaacctg tacaagttaa ctgctagtgc ctacacataca 2520
 gagagaaaaac tgaaacaggc aataaactat tcaaaaagtt tagacatgga gaaaggagtt 2580
 15 gagaatgacc taagctatca gaatatagga ggaaacacca attggaagcc tccgctcaac 2640
 tgtaagattt ataactacct gaacaggatt ggatgcttct tccctcatcc tcgctgtagt 2700
 aaaagaaaag atgctgctga ttttgccata tgtatgcatg ctggccgtct agatgaacaa 2760
 ctaccaagc aaattccttt caccatcctc tcaggagatc aaggttttct ggagctagag 2820
 20 aatcaattta agaagactca gaggccagct catatactaa accctcacca cttagaggga 2880
 gatatgatgt gtgccttggt aaatagcata tctgatacca ccaaagaatg tgacagtgat 2940
 gataacatgg gtgccaaaaa tacttcaata ggagaagaat ttatatccac agaagatgtg 3000
 25 gaattagaag aagctattag aagaagtctt gaggaatgt aa 3042

<210> 6

30 <211> 3042

<212> ADN

<213> *Homo sapiens*

35 <400> 6

atgggagacc cggggtcgga gataatagaa tctgtccctc cagctggccc tgaggcatct 60
 40 gagtcaacaa cggatgaaaa tgaagacgac attcagtttg tcagtgaagg accattacga 120
 cctgttcttg aatacattga tctggtcagc agtgatgatg aagagcctag cacctcttat 180
 actgatgaga atattaaacg taaagaccat attgattatc agaaggataa agttgcttta 240
 actctggctc gtctagcccg ccatgttgaa gtggagaaac agcagaaaga agagaagaat 300
 45 agagcattca gagaaaaaat tgattttcag catgctcatg ggttacaaga attggaattt 360
 attcgaggac attctgatac agaagcagca agactgtgtg tggaccagtg gctaaaaatg 420
 ccaggactca aaacaggcac aattaattgt ggaacaaaaa gttcattccg aagaggaggc 480
 cacacgtggg tgtctgggaa accaacttta tgtcctataa tgcaactgaa caaggagttt 540
 50 gacaatgggc accttctctt aggacatttg aaaaggttcg atcactctcc atgtgatcca 600
 acaattacac tacatggacc tttcttcagc tcccttgctt gtgtagtatg ttataaaaaa 660
 tttgttactc aacaacaata tagagatcac ctttttgata aggaagccac agatgatgga 720
 cataacaaca accttcttcc tcagattatt cagtgttttg catgtccaaa ttgcttcctt 780
 55 ctttttagca gaaaggagga gtgttcaaag catatgtctg gaaagaatca tttccatcag 840
 agtttcaaac tgggtgataa caaaggaatt gcacatccaa tatctttccc atcttttgca 900
 aagaaacttt tgatctctct gtgcaaagat gttcccttcc aagttaagtg tgtggcctgc 960
 60 cacaagacac tgcgttccca catggagctc actgccatt tcagagttca ttgtcgaaat 1020
 gctggacctg tagctgtagc tgagaagagc attaccaggg ttgcagagaa attcatatta 1080
 agaggttatt gtccagattg caatcaagtc tttgtggatg aaaccagcac ccaaaatcat 1140
 aagcagaatt caggacacaa agtccgagtc attaaactag tggagaatc agtcttactc 1200
 65 tattgccaca gcagcgaagg gaacaaggat ccttcttctg acttgcatth attgttggat 1260
 caatcaaat tttcatcact taaaagaacc atgtctatta aagaatctag ctactggag 1320

ES 2 305 116 T3

```

5  tgcattgccca ttccaaaaaa gaagatgaat ttaaaagata aaagccatga aggtgttgct 1380
   tgtgtccaga aagaaaaatc agtagttaa acctggttct gtgaatgcaa tcagcgattc 1440
10 ccaagtgaag atgcagtaga aaagcatgtt ttctcagcaa acacaatggg ttataaatgt 1500
   gtggtctgtg gaaaggtatg tgatgattca ggggtcattc gtttacacat gagccggatt 1560
   cacggagggg cacatttaaa taactttctt ttctgggtgc ggacatgcaa aaaggagtta 1620
   acaaggaaaag atactatcat ggcacatgtg actgaatttc ataatggaca cagatatttt 1680
15 tatgagatgg atgaggtaga aggtgaaact ttgccatcat cctctacaac attggataat 1740
   ttgactgcta acaagccttc atcagctatt actgttattg atcattcccc ggcaaatagt 1800
   tctccgaggg gtaaatggca atgccggatt tgtgaagata tgtttgattc ccaggaatat 1860
   gtaaaacagc actgcatgtc tttggcaagc cacaagtttc atagatacag ctgtgctcac 1920
15 tgcagaaaagc cttttcataa gatagaaaca ttgtaccgac attgocaaga tgagcatgac 1980
   aatgagataa agattaaata cttctgtggg ctttgtgac tttatcttta tgtggaagaa 2040
   gcattttctga gtcattatga ggagcaccac agcatagatt atgtatttgt gtcagaaaaa 2100
20 actgaaaactt caattaaaac cgaagatgat tttccagtaa tagagaccag taaccagtta 2160
   acttgtggtt gccgtgagag ttacatctgt aaagtcaaca gaaaagaaga ttatagcaga 2220
   tgtctcmeta tcatgctgga taaaggaaaa ctgtggtttc gctgcagttt atgttcggca 2280
   acagcacaga atttaaccga catgaacact catatccatc aagtgcacaa agaaaagagt 2340
25 gatgaggagg agcagcagta tgtaatcaag tgtggcacct gcaccaaagc atttcatgat 2400
   cctgagagtg cacagcagca tttccataga aaacattgct tottacagaa acccagtggtg 2460
   gctcattttg gatctgmeta atcaaacctg tacaagttta ctgctagtgc ctcacataca 2520
   gagagaaaac tgaaacaggc aataaactat tcaaaaagtt tagacatgga gaaaggagtt 2580
30 gagaatgacc taagctatca gaatatagga ggaaacacca attggaagcc tccgctcaac 2640
   tgtaagattt ataactacct gaacaggatt ggatgcttct tccctcatcc tccgctgtagt 2700
   aaaagaaaag atgctgctga ttttgccata tgtatgcatg ctggcogtct agatgaacaa 2760
   ctacccaagc aaattccttt caccatcctc tcaggagatc aaggttttct ggagctagag 2820
35 aatcaattta agaagactca gaggccagct catatactaa accctcacca cttagagggg 2880
   gatatgatgt gtgccttggt aaatagcata tctgatacca ccaaagaatg tgacagtgat 2940
   gataacatgg gtgccaaaaa tacttcaata ggagaagaat ttatatccac agaagatgtg 3000
40 gaattagaag aagctattag aagaagtctt gaggaatgt aa 3042

```

<210> 7

<211> 1013

<212> PRT

<213> *Homo sapiens*

<400> 7

```

55  Met Gly Asp Pro Gly Ser Glu Ile Ile Glu Ser Val Pro Pro Ala Gly
     1             5             10             15

     Pro Glu Ala Ser Glu Ser Thr Thr Asp Glu Asn Glu Asp Asp Ile Gln
           20             25             30

60  Phe Val Ser Glu Gly Pro Leu Arg Pro Val Leu Glu Tyr Ile Asp Leu
           35             40             45

65  Val Ser Ser Asp Asp Glu Glu Pro Ser Thr Ser Tyr Thr Asp Glu Asn
           50             55             60

```

ES 2 305 116 T3

Ile Lys Arg Lys Asp His Ile Asp Tyr Gln Lys Asp Lys Val Ala Leu
 65 70 75 80
 Thr Leu Ala Arg Leu Ala Arg His Val Glu Val Glu Lys Gln Gln Lys
 85 90 95
 Glu Glu Lys Asn Arg Ala Phe Arg Glu Lys Ile Asp Phe Gln His Ala
 100 105 110
 His Gly Leu Gln Glu Leu Glu Phe Ile Arg Gly His Ser Asp Thr Glu
 115 120 125
 Ala Ala Arg Leu Cys Val Asp Gln Trp Leu Lys Met Pro Gly Leu Lys
 130 135 140
 Thr Gly Thr Ile Asn Cys Gly Thr Lys Ser Ser Phe Arg Arg Gly Gly
 145 150 155 160
 His Thr Trp Val Ser Gly Lys Pro Ile Leu Cys Pro Ile Met His Cys
 165 170 175
 Asn Lys Glu Phe Asp Asn Gly His Leu Leu Leu Gly His Leu Lys Arg
 180 185 190
 Phe Asp His Ser Pro Cys Asp Pro Thr Ile Thr Leu His Gly Pro Phe
 195 200 205
 Phe Ser Ser Phe Ala Cys Val Val Cys Tyr Lys Lys Phe Val Thr Gln
 210 215 220
 Gln Gln Tyr Arg Asp His Leu Phe Asp Lys Glu Ala Thr Asp Asp Gly
 225 230 235 240
 His Asn Asn Asn Leu Leu Pro Gln Ile Ile Gln Cys Phe Ala Cys Pro
 245 250 255
 Asn Cys Phe Leu Leu Phe Ser Arg Lys Glu Glu Cys Ser Lys His Met
 260 265 270
 Ser Gly Lys Asn His Phe His Gln Ser Phe Lys Leu Gly Asp Asn Lys
 275 280 285
 Gly Ile Ala His Pro Ile Ser Phe Pro Ser Phe Ala Lys Lys Leu Leu
 290 295 300
 Ile Ser Leu Cys Lys Asp Val Pro Phe Gln Val Lys Cys Val Ala Cys
 305 310 315 320

ES 2 305 116 T3

5 His Lys Thr Leu Arg Ser His Met Glu Leu Thr Ala His Phe Arg Val
 325 330 335
 His Cys Arg Asn Ala Gly Pro Val Ala Val Ala Glu Lys Ser Ile Thr
 340 345 350
 10 Gln Val Ala Glu Lys Phe Ile Leu Arg Gly Tyr Cys Pro Asp Cys Asn
 355 360 365
 15 Gln Val Phe Val Asp Glu Thr Ser Thr Gln Asn His Lys Gln Asn Ser
 370 375 380
 Gly His Lys Val Arg Val Ile Asn Ser Val Glu Glu Ser Val Leu Leu
 385 390 395 400
 20 Tyr Cys His Ser Ser Glu Gly Asn Lys Asp Pro Ser Ser Asp Leu His
 405 410 415
 25 Leu Leu Leu Asp Gln Ser Lys Phe Ser Ser Leu Lys Arg Thr Met Ser
 420 425 430
 30 Ile Lys Glu Ser Ser Ser Leu Glu Cys Ile Ala Ile Pro Lys Lys Lys
 435 440 445
 35 Met Asn Leu Lys Asp Lys Ser His Glu Gly Val Ala Cys Val Gln Lys
 450 455 460
 40 Glu Lys Ser Val Val Lys Thr Trp Phe Cys Glu Cys Asn Gln Arg Phe
 465 470 475 480
 Pro Ser Glu Asp Ala Val Glu Lys His Val Phe Ser Ala Asn Thr Met
 485 490 495
 45 Gly Tyr Lys Cys Val Val Cys Gly Lys Val Cys Asp Asp Ser Gly Val
 500 505 510
 50 Ile Arg Leu His Met Ser Arg Ile His Gly Gly Ala His Leu Asn Asn
 515 520 525
 Phe Leu Phe Trp Cys Arg Thr Cys Lys Lys Glu Leu Thr Arg Lys Asp
 530 535 540
 55 Thr Ile Met Ala His Val Thr Glu Phe His Asn Gly His Arg Tyr Phe
 545 550 555 560
 60 Tyr Glu Met Asp Glu Val Glu Gly Glu Thr Leu Pro Ser Ser Ser Thr
 565 570 575
 65

ES 2 305 116 T3

Thr Leu Asp Asn Leu Thr Ala Asn Lys Pro Ser Ser Ala Ile Thr Val
 580 585 590
 5
 Ile Asp His Ser Pro Ala Asn Ser Ser Pro Arg Gly Lys Trp Gln Cys
 595 600 605
 10
 Arg Ile Cys Glu Asp Met Phe Asp Ser Gln Glu Tyr Val Lys Gln His
 610 615 620
 15
 Cys Met Ser Leu Ala Ser His Lys Phe His Arg Tyr Ser Cys Ala His
 625 630 635 640
 20
 Cys Arg Lys Pro Phe His Lys Ile Glu Thr Leu Tyr Arg His Cys Gln
 645 650 655
 25
 Asp Glu His Asp Asn Glu Ile Lys Ile Lys Tyr Phe Cys Gly Leu Cys
 660 665 670
 30
 Asp Leu Ile Phe Asn Val Glu Glu Ala Phe Leu Ser His Tyr Glu Glu
 675 680 685
 35
 His His Ser Ile Asp Tyr Val Phe Val Ser Glu Lys Thr Glu Thr Ser
 690 695 700
 40
 Ile Lys Thr Glu Asp Asp Phe Pro Val Ile Glu Thr Ser Asn Gln Leu
 705 710 715 720
 45
 Thr Cys Gly Cys Arg Glu Ser Tyr Ile Cys Lys Val Asn Arg Lys Glu
 725 730 735
 50
 Asp Tyr Ser Arg Cys Leu Gln Ile Met Leu Asp Lys Gly Lys Leu Trp
 740 745 750
 55
 Phe Arg Cys Ser Leu Cys Ser Ala Thr Ala Gln Asn Leu Thr Asp Met
 755 760 765
 60
 Asn Thr His Ile His Gln Val His Lys Glu Lys Ser Asp Glu Glu Glu
 770 775 780
 65
 Gln Gln Tyr Val Ile Lys Cys Gly Thr Cys Thr Lys Ala Phe His Asp
 785 790 795 800
 Pro Glu Ser Ala Gln Gln His Phe His Arg Lys His Cys Phe Leu Gln
 805 810 815
 Lys Pro Ser Val Ala His Phe Gly Ser Glu Lys Ser Asn Leu Tyr Lys
 820 825 830

ES 2 305 116 T3

Phe Thr Ala Ser Ala Ser His Thr Glu Arg Lys Leu Lys Gln Ala Ile
 835 840 845
 5 Asn Tyr Ser Lys Ser Leu Asp Met Glu Lys Gly Val Glu Asn Asp Leu
 850 855 860
 10 Ser Tyr Gln Asn Ile Gly Gly Asn Thr Asn Trp Lys Pro Pro Leu Asn
 865 870 875 880
 15 Cys Lys Ile Tyr Asn Tyr Leu Asn Arg Ile Gly Cys Phe Phe Leu His
 885 890 895
 20 Pro Arg Cys Ser Lys Arg Lys Asp Ala Ala Asp Phe Ala Ile Cys Met
 900 905 910
 25 His Ala Gly Arg Leu Asp Glu Gln Leu Pro Lys Gln Ile Pro Phe Thr
 915 920 925
 30 Ile Leu Ser Gly Asp Gln Gly Phe Leu Glu Leu Glu Asn Gln Phe Lys
 930 935 940
 35 Lys Thr Gln Arg Pro Ala His Ile Leu Asn Pro His His Leu Glu Gly
 945 950 955 960
 40 Asp Met Met Cys Ala Leu Leu Asn Ser Ile Ser Asp Thr Thr Lys Glu
 965 970 975
 45 Cys Asp Ser Asp Asp Asn Met Gly Ala Lys Asn Thr Ser Ile Gly Glu
 980 985 990
 50 Glu Phe Ile Ser Thr Glu Asp Val Glu Leu Glu Glu Ala Ile Arg Arg
 995 1000 1005
 55 Ser Leu Glu Glu Met
 1010
 60 Met Gly Asp Pro Gly Ser Glu Ile Ile Glu Ser Val Pro Pro Ala Gly
 1 5 10 15
 65 Pro Glu Ala Ser Glu Ser Thr Thr Asp Glu Asn Glu Asp Asp Ile Gln

ES 2 305 116 T3

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

ES 2 305 116 T3

	275		280		285
5	Gly Ile Ala His Pro Ile Ser Phe Pro Ser Phe Ala Lys Lys Leu Leu				
	290		295		300
10	Ile Ser Leu Cys Lys Asp Val Pro Phe Gln Val Lys Cys Val Ala Cys				
	305		310		315 320
15	His Lys Thr Leu Arg Ser His Met Glu Leu Thr Ala His Phe Arg Val				
		325		330	335
20	His Cys Arg Asn Ala Gly Pro Val Ala Val Ala Glu Lys Ser Ile Thr				
		340		345	350
25	Gln Val Ala Glu Lys Phe Ile Leu Arg Gly Tyr Cys Pro Asp Cys Asn				
		355		360	365
30	Gln Val Phe Val Asp Glu Thr Ser Thr Gln Asn His Lys Gln Asn Ser				
		370		375	380
35	Gly His Lys Val Arg Val Ile Asn Ser Val Glu Glu Ser Val Leu Leu				
		385		390	395 400
40	Tyr Cys His Ser Ser Glu Gly Asn Lys Asp Pro Ser Ser Asp Leu His				
		405		410	415
45	Leu Leu Leu Asp Gln Ser Lys Phe Ser Ser Leu Lys Arg Thr Met Ser				
		420		425	430
50	Ile Lys Glu Ser Ser Ser Leu Glu Cys Ile Ala Ile Pro Lys Lys Lys				
		435		440	445
55	Met Asn Leu Lys Asp Lys Ser His Glu Gly Val Ala Cys Val Gln Lys				
		450		455	460
60	Glu Lys Ser Val Val Lys Thr Trp Phe Cys Glu Cys Asn Gln Arg Phe				
		465		470	475 480
65	Pro Ser Glu Asp Ala Val Glu Lys His Val Phe Ser Ala Asn Thr Met				
		485		490	495
70	Gly Tyr Lys Cys Val Val Cys Gly Lys Val Cys Asp Asp Ser Gly Val				
		500		505	510
75	Ile Arg Leu His Met Ser Arg Ile His Gly Gly Ala His Leu Asn Asn				
		515		520	525
80	Phe Leu Phe Trp Cys Arg Thr Cys Lys Lys Glu Leu Thr Arg Lys Asp				

ES 2 305 116 T3

	530		535		540														
5	Thr	Ile	Met	Ala	His	Val	Thr	Glu	Phe	His	Asn	Gly	His	Arg	Tyr	Phe			
	545					550					555					560			
10	Tyr	Glu	Met	Asp	Glu	Val	Glu	Gly	Glu	Thr	Leu	Pro	Ser	Ser	Ser	Thr			
					565					570						575			
15	Thr	Leu	Asp	Asn	Leu	Thr	Ala	Asn	Lys	Pro	Ser	Ser	Ala	Ile	Thr	Val			
				580					585					590					
20	Ile	Asp	His	Ser	Pro	Ala	Asn	Ser	Ser	Pro	Arg	Gly	Lys	Trp	Gln	Cys			
		595						600					605						
25	Arg	Ile	Cys	Glu	Asp	Met	Phe	Asp	Ser	Gln	Glu	Tyr	Val	Lys	Gln	His			
	610						615					620							
30	Cys	Met	Ser	Leu	Ala	Ser	His	Lys	Phe	His	Arg	Tyr	Ser	Cys	Ala	His			
	625					630					635					640			
35	Cys	Arg	Lys	Pro	Phe	His	Lys	Ile	Glu	Thr	Leu	Tyr	Arg	His	Cys	Gln			
				645					650						655				
40	Asp	Glu	His	Asp	Asn	Glu	Ile	Lys	Ile	Lys	Tyr	Phe	Cys	Gly	Leu	Cys			
				660					665					670					
45	Asp	Leu	Ile	Phe	Asn	Val	Glu	Glu	Ala	Phe	Leu	Ser	His	Tyr	Glu	Glu			
		675					680						685						
50	His	His	Ser	Ile	Asp	Tyr	Val	Phe	Val	Ser	Glu	Lys	Thr	Glu	Thr	Ser			
	690						695					700							
55	Ile	Lys	Thr	Glu	Asp	Asp	Phe	Pro	Val	Ile	Glu	Thr	Ser	Asn	Gln	Leu			
	705					710					715					720			
60	Thr	Cys	Gly	Cys	Arg	Glu	Ser	Tyr	Ile	Cys	Lys	Val	Asn	Arg	Lys	Glu			
				725					730						735				
65	Asp	Tyr	Ser	Arg	Cys	Leu	Gln	Ile	Met	Leu	Asp	Lys	Gly	Lys	Leu	Trp			
			740					745					750						
70	Phe	Arg	Cys	Ser	Leu	Cys	Ser	Ala	Thr	Ala	Gln	Asn	Leu	Thr	Asp	Met			
		755						760					765						
75	Asn	Thr	His	Ile	His	Gln	Val	His	Lys	Glu	Lys	Ser	Asp	Glu	Glu	Glu			
	770					775						780							
80	Gln	Gln	Tyr	Val	Ile	Lys	Cys	Gly	Thr	Cys	Thr	Lys	Ala	Phe	His	Asp			

ES 2 305 116 T3

	785		790		795		800									
5	Pro	Glu	Ser	Ala	Gln	Gln	His	Phe	His	Arg	Lys	His	Cys	Phe	Leu	Gln
				805						810					815	
10	Lys	Pro	Ser	Val	Ala	His	Phe	Gly	Ser	Glu	Lys	Ser	Asn	Leu	Tyr	Lys
				820					825					830		
15	Phe	Thr	Ala	Ser	Ala	Ser	His	Thr	Glu	Arg	Lys	Leu	Lys	Gln	Ala	Ile
			835					840					845			
20	Asn	Tyr	Ser	Lys	Ser	Leu	Asp	Met	Glu	Lys	Gly	Val	Glu	Asn	Asp	Leu
		850					855					860				
25	Ser	Tyr	Gln	Asn	Ile	Gly	Gly	Asn	Thr	Asn	Trp	Lys	Pro	Pro	Leu	Asn
	865				870					875					880	
30	Cys	Lys	Ile	Tyr	Asn	Tyr	Leu	Asn	Arg	Ile	Gly	Cys	Phe	Phe	Leu	His
				885					890						895	
35	Pro	Arg	Cys	Ser	Lys	Arg	Lys	Asp	Ala	Ala	Asp	Phe	Ala	Ile	Cys	Met
			900						905					910		
40	His	Ala	Gly	Arg	Leu	Asp	Glu	Gln	Leu	Pro	Lys	Gln	Ile	Pro	Phe	Thr
		915					920					925				
45	Ile	Leu	Ser	Gly	Asp	Gln	Gly	Phe	Leu	Glu	Leu	Glu	Asn	Gln	Phe	Lys
	930					935						940				
50	Lys	Thr	Gln	Arg	Pro	Ala	His	Ile	Leu	Asn	Pro	His	His	Leu	Glu	Gly
	945				950					955					960	
55	Asp	Met	Met	Cys	Ala	Leu	Leu	Asn	Ser	Ile	Ser	Asp	Thr	Thr	Lys	Glu
				965					970						975	
60	Cys	Asp	Ser	Asp	Asp	Asn	Met	Gly	Ala	Lys	Asn	Thr	Ser	Ile	Gly	Glu
			980					985						990		
65	Glu	Phe	Ile	Ser	Thr	Glu	Asp	Val	Glu	Leu	Glu	Glu	Ala	Ile	Arg	Arg
		995				1000						1005				
70	Ser	Leu	Glu	Glu	Met											
	1010															
75	<210>	9														
80	<211>	3183														
85	<212>	ADN														
90	<213>	Homo sapiens														

ES 2 305 116 T3

<400> 9

```

5  atgggagacc cggggtcgga gataatagaa tctgtccctc cagctggccc .tgaggcatct 60
   gagtcaacaa cggatgaaaa tgaagacgac attcagtttg tcagtgaagg accattacga 120
   cctgttcttg aatacattga tctgggtcagc agtgatgatg aagagcctag cacctcttat 180
   actgatgaga atattaaacg taaagaccat attgattatc agaaggataa agttgcttta 240
10  actctggctc gtctagcccg ccatgttgaa gtggagaaac agcagaaaga agagaagaat 300
   agagcattca gaaaaaaat tgattttcag catgctcatg gggtacaaga attggaattt 360
   attcgaggac attctgatac agaagcagca agactgtgtg tggaccagtg gctaaaaatg 420
   ccaggactca aaacaggcac aattaattgt ggaacaaaaa gttcattccg aagaggaggc 480
15  cacacgtggg tgtctgggaa accaatttta tgtcctataa tgcactgtaa caaggagttt 540
   gacaatgggc accttctctt aggacatttg aaaaggttcg atcactctcc atgtgatcca 600
   acaattacac tacatggacc tttcttcagc tcctttgctt gtgtagtatg ttataaaaaa 660
   tttgttactc aacaacaata tagagatcac ctttttgata aggaagccac agatgatgga 720
20  cataacaaca accttcttcc tcagattatt cagtgttttg catgtccaaa ttgcttcctt 780
   ctttttagca gaaaggagga gtgttcaaag catatgtctg gaaagaatca tttccatcag 840
   agtttcaaac tgggtgataa caaaggaatt gcacatccaa tatctttccc atcttttgca 900
   aagaaacttt tgatctctct gtgcaaagat gttccctttc aagttaagtg tgtggcctgc 960
25  cacaagacac tgcgttccca catggagctc actgccatt tcagagttca ttgtcgaaat 1020
   gctggacctg tagctgtagc tgagaagagc attaccocagg ttgcagagaa attcatatta 1080
   agaggttatt gtccagattg caatcaagtc tttgtggatg aaaccagcac ccaaaatcat 1140
   aagcagaatt caggacacaa agtccgagtc attaactcag tggagaatc agtcttactc 1200
30  tattgccaca gcagcgaagg gaacaaggat ccttcttctg acttgcatth attgttggat 1260
   caatcaaaat tttcatcact taaaagaacc atgtctatta aagaatctag ctactggag 1320
   tgcattgcca ttccaaaaaa gaagatgaat taaaagata aaagccatga aggtgttgc 1380
   tgtgtccaga aagaaaaatc agtagttaaa acctggttct gtgaatgcaa tcagcgattc 1440
35  ccaagtgaag atgcagtaga aaagcatgtt ttctcagcaa acacaatggg ttataaatgt 1500
   gtggtctgtg gaaaggtagt tgatgattca ggggtcattc gtttacacat gagccggatt 1560
   cacggagggg cacatttaaa taactttctt ttctgggtgc ggacatgcaa aaaggagtta 1620
   acaaggaaag atactatcat ggcacatgtg actgaatttc ataattggaca cagatatttt 1680
40  tatgagatgg atgaggtaga aggtgaaact ttgccatcat cctctacaac attggataat 1740
   ttgactgcta acaagccttc atcagctatt actgttattg atcattcccc ggcaaatagt 1800
   tctccgaggg gtaaattggca atgcoggatt tgtgaagata tgtttgattc ccaggaatat 1860
   gtaaaacagc actgcatgtc tttggcaagc cacaagtttc atagatacag ctgtgctcac 1920
45  tgcagaaagc cttttcataa gatagaaaca ttgtaccgac attgccaaga tgagcatgac 1980
   aatgagataa agattaaata cttctgtggg ctttgtgatc ttatctttaa tgtggaagaa 2040
   gcatttctga gtcattatga ggagcaaccac agcatagatt atgtatttgt gtcagaaaaa 2100
   actgaaactt caattaaaac ogaagatgat tttccagtaa tagagaccag taaccagtta 2160
50  acttgtggtt gccgtgagag ttacatctgt aaagtcaaca gaaaagaaga ttatagcaga 2220
   tgtctccaaa tcatgctgga taaaggaaaa ctgtggtttc gctgcagttt atgttcggca 2280
   acagcacaga atttaaccga catgaacact catatccatc aagtgcacaa agaaaagagt 2340
55  gatgaggagg agcagcagta tgtaatcaag tgtggcacct gcaccaaagc atttcatgat 2400
   cctgagagtg cacagcagca tttccataga aaacattgct tcttacagaa acccagtgtg 2460
   gtcatttttg gatctgaaaa atcaaacctg tacaagttta ctgctagtgc ctacataca 2520
   gagagaaaac tgaaacaggc aataaactat tcaaaaagtt tagacatgga gaaaggagtt 2580
60  gagaatgacc taagctatca gaatatagag gaagaaattg ttgagcttcc agatttggat 2640
   tacctgcgaa ccatgactca tatagtcttt gtagattttg ataactggtc aaactttttt 2700

```

65

ES 2 305 116 T3

```

gggtcatctac cagggcatct aaaccaagga acatttattt ggggctttca aggaggaaac 2760
accaattgga agcctccgct caactgtaag atttataact acctgaacag gattggatgc 2820
5 ttcttctctc atcctcgcgt tagtaaaaga aaagatgctg ctgattttgc catatgtatg 2880
catgctggcc gtctagatga acaactaccc aagcaaattc ctttcacccat cctctcagga 2940
gatcaagggtt ttctggagct agagaatcaa tttaagaaga ctcagaggcc agctcatata 3000
ctaaaccctc accacttaga gggagatatg atgtgtgcct tgttaaatac catatctgat 3060
10 accaccaag aatgtgacag tgatgataac atgggtgcca aaaatacttc aataggagaa 3120
gaatttatat ccacagaaga tgtggaatta gaagaagcta ttagaagaag tcttgaggaa 3180
atg 3183

```

<210> 10

<211> 3183

<212> ADN

<213> *Homo sapiens*

<400> 10

```

atgggagacc cggggtcggg gataatagaa tctgtccctc cagctggccc tgaggcatct 60
gagtcaacaa cggatgaaaa tgaagacgac attcagtttg tcagtgaagg accattacga 120
30 cctgttcttg aatacattga tctggtcagc agtgatgatg aagagcctag cacctcttat 180
actgatgaga atattaaacg taaagaccat attgattatc agaaggataa agttgcttta 240
actctggctc gtctagcccg ccatgttgaa gtggagaaac agcagaaaga agagaagaat 300
agagcattca gagaaaaaat tgattttcag catgctcatg gggtacaaga attggaattt 360
35 attcaggac attctgatac agaagcagca agactgtgtg tggaccagtg gctaaaaatg 420
ccaggactca aaacaggcac aattaattgt ggaacaaaaa gtccattccg aagaggaggc 480
cacacgtggg tgtctgggaa accaacttta tgtcctataa tgcactgtaa caaggagttt 540
gacaatgggc accttctctt aggcattttg aaaaggttcg atcactctcc atgtgatoca 600
40 acaattacac tacatggacc tttcttcagc tcttttgctt gtgtagtatg ttataaaaaa 660
tttgttactc aaacaaatac tagagatcac ctttttgata aggaagccac agatgatgga 720
cataacaaca accttcttcc tcagattatt cagtgttttg catgtccaaa ttgcttctct 780
cttttttagc gaaaggagga gtgttcaaag catatgtctg gaaagaatca tttccatcag 840
45 agtttcaaac tgggtgataa caaaggaatt gcacatccaa tatctttccc atcttttgca 900
aagaaacttt tgatctctct gtgcaaagat gttccctttc aagttaagtg tgtggcctgc 960
cacaagacac tgcgttccca catggagctc actgccatt tcagagttca ttgtcgaaat 1020
getggacctg tagctgtagc tgagaagagc attaccagg ttgcagagaa attcatatta 1080
50 agaggttatt gtccagattg caatcaagtc tttgtggatg aaaccagcac ccaaaatcat 1140
aagcagaatt caggacacaa agtccgagtc attaactcag tggaaagaatc agtcttactc 1200
tattgccaca gcagcgaagg gaacaaggat ccttcttctg acttgcatth attgttggtt 1260
caatcaaaat tttcatcact taaaagaacc atgtctatta aagaatctag ctactggag 1320
55 tgcattgcca ttccaaaaaa gaagatgaat ttaaaagata aaagccatga aggtgttgct 1380
tgtgtccaga aagaaaaatc agtagttaa acctggttct gtgaatgcaa tcagcgattc 1440
ccaagtgaag atgcagtaga aaagcatgtt ttctcagcaa acacaatggg ttataaatgt 1500
60 gtggtctgtg gaaaggtagt tgatgattca ggggtcattc gtttacacat gagccggatt 1560
cacggagggg cacatttaaa taactttctt ttctgtgtgc ggacatgcaa aaaggagtta 1620
acaaggaaag atactatcat ggcacatgtg actgaatttc ataattggaca cagatatattt 1680
tatgagatgg atgaggtaga aggtgaaact ttgccatcat cctctacaac attggataat 1740
65 ttgactgcta acaagccttc atcagctatt actgttattg atcattcccc ggcaaatagt 1800
tctccgaggg gtaaatggca atgccggatt tgtgaagata tgtttgattc ccagggaatat 1860

```

ES 2 305 116 T3

```

gtaaaacagc actgcatgtc tttggcaagc cacaagtttc atagatacag ctgtgctcac 1920
tgcagaaagc cttttcataa gatagaaaca ttgtaccgac attgccaaga tgagcatgac 1980
5 aatgagataa agattaaata cttctgtggg ctttgtgatc ttatctttta tgtggaagaa 2040
gcatttctga gtcattatga ggagcaccac agcatagatt atgtatttgt gtcagaaaaa 2100
actgaaactt caattaaaac cgaagatgat tttccagtaa tagagaccag taaccagtta 2160
acttgtgggt gccgtgagag ttacatctgt aaagtcaaca gaaaagaaga ttatagcaga 2220
10 tgtctccaaa tcatgctgga taaaggaaaa ctgtggtttc gctgcagttt atgttcggca 2280
acagcacaga atttaaccga catgaacact catatccatc aagtgcacaa agaaaagagt 2340
gatgaggagg agcagcagta tgtaatcaag tgtggcacct gcaccaaagc atttcatgat 2400
cctgagagtg cacagcagca tttccataga aaacattgct tcttacagaa acccagtgtg 2460
15 gctcatittg gatctgaaaa atcaaaccctg tacaagttaa ctgctagtgc ctcacatata 2520
gagagaaaaac tgaaacaggc aataaactat tcaaaaagtt tagacatgga gaaaggagtt 2580
gagaatgacc taagctatca gaatatagag gaagaaattg ttgagcttcc agatttggat 2640
tacctgcgaa ccatgactca tatagtcttt gtagattttg ataactggtc aaactttttt 2700
20 ggtcatctac cagggcactt aaaccaagga acattttatt ggggctttca aggaggaaac 2760
accaattgga agcctccgct caactgtaag atttataact acctgaacag gattggatgc 2820
ttcttccttc atcctcgctg tagtaaaaga aaagatgctg ctgattttgc catatgtatg 2880
25 catgctggcc gtctagatga acaactaccc aagcaaattc ctttcaccat cctctcagga 2940
gatcaagggt ttgtggagct agagaatcaa tttaagaaga ctcagaggcc agctcatata 3000
ctaaaccctc accacttaga gggagatatg atgtgtgcct tgttaaatag catatctgat 3060
accaccaaag aatgtgacag tgatgataac atgggtgcc aaaaacttc aataggagaa 3120
30 gaatttatat ccacagaaga tgtggaatta gaagaagcta ttagaagaag tcttgaggaa 3180
atg 3183

```

```

35 <210> 11
    <211> 1061
    <212> PRT
    <213> Homo sapiens
40 <400> 11

```

```

45 Met Gly Asp Pro Gly Ser Glu Ile Ile Glu Ser Val Pro Pro Ala Gly
    1 5 10 15

Pro Glu Ala Ser Glu Ser Thr Thr Asp Glu Asn Glu Asp Asp Ile Gln
50 20 25 30

Phe Val Ser Glu Gly Pro Leu Arg Pro Val Leu Glu Tyr Ile Asp Leu
55 35 40 45

Val Ser Ser Asp Asp Glu Glu Pro Ser Thr Ser Tyr Thr Asp Glu Asn
60 50 55 60

Ile Lys Arg Lys Asp His Ile Asp Tyr Gln Lys Asp Lys Val Ala Leu
65 65 70 75 80

Thr Leu Ala Arg Leu Ala Arg His Val Glu Val Glu Lys Gln Gln Lys
65 85 90 95

```

ES 2 305 116 T3

Glu Glu Lys Asn Arg Ala Phe Arg Glu Lys Ile Asp Phe Gln His Ala
 100 105 110
 5
 His Gly Leu Gln Glu Leu Glu Phe Ile Arg Gly His Ser Asp Thr Glu
 115 120 125
 10
 Ala Ala Arg Leu Cys Val Asp Gln Trp Leu Lys Met Pro Gly Leu Lys
 130 135 140
 15
 Thr Gly Thr Ile Asn Cys Gly Thr Lys Ser Ser Phe Arg Arg Gly Gly
 145 150 155 160
 20
 His Thr Trp Val Ser Gly Lys Pro Ile Leu Cys Pro Ile Met His Cys
 165 170 175
 25
 Asn Lys Glu Phe Asp Asn Gly His Leu Leu Leu Gly His Leu Lys Arg
 180 185 190
 30
 Phe Asp His Ser Pro Cys Asp Pro Thr Ile Thr Leu His Gly Pro Phe
 195 200 205
 35
 Phe Ser Ser Phe Ala Cys Val Val Cys Tyr Lys Lys Phe Val Thr Gln
 210 215 220
 40
 Gln Gln Tyr Arg Asp His Leu Phe Asp Lys Glu Ala Thr Asp Asp Gly
 225 230 235 240
 45
 His Asn Asn Asn Leu Leu Pro Gln Ile Ile Gln Cys Phe Ala Cys Pro
 245 250 255
 50
 Asn Cys Phe Leu Leu Phe Ser Arg Lys Glu Glu Cys Ser Lys His Met
 260 265 270
 55
 Ser Gly Lys Asn His Phe His Gln Ser Phe Lys Leu Gly Asp Asn Lys
 275 280 285
 60
 Gly Ile Ala His Pro Ile Ser Phe Pro Ser Phe Ala Lys Lys Leu Leu
 290 295 300
 65
 Ile Ser Leu Cys Lys Asp Val Pro Phe Gln Val Lys Cys Val Ala Cys
 305 310 315 320
 His Lys Thr Leu Arg Ser His Met Glu Leu Thr Ala His Phe Arg Val
 325 330 335
 His Cys Arg Asn Ala Gly Pro Val Ala Val Ala Glu Lys Ser Ile Thr
 340 345 350

ES 2 305 116 T3

5 Gln Val Ala Glu Lys Phe Ile Leu Arg Gly Tyr Cys Pro Asp Cys Asn
 355 360 365
 10 Gln Val Phe Val Asp Glu Thr Ser Thr Gln Asn His Lys Gln Asn Ser
 370 375 380
 15 Tyr Cys His Ser Ser Glu Gly Asn Lys Asp Pro Ser Ser Asp Leu His
 405 410 415
 20 Leu Leu Leu Asp Gln Ser Lys Phe Ser Ser Leu Lys Arg Thr Met Ser
 420 425 430
 25 Ile Lys Glu Ser Ser Ser Leu Glu Cys Ile Ala Ile Pro Lys Lys Lys
 435 440 445
 30 Met Asn Leu Lys Asp Lys Ser His Glu Gly Val Ala Cys Val Gln Lys
 450 455 460
 35 Glu Lys Ser Val Val Lys Thr Trp Phe Cys Glu Cys Asn Gln Arg Phe
 465 470 475 480
 40 Pro Ser Glu Asp Ala Val Glu Lys His Val Phe Ser Ala Asn Thr Met
 485 490 495
 45 Gly Tyr Lys Cys Val Val Cys Gly Lys Val Cys Asp Asp Ser Gly Val
 500 505 510
 50 Ile Arg Leu His Met Ser Arg Ile His Gly Gly Ala His Leu Asn Asn
 515 520 525
 55 Phe Leu Phe Trp Cys Arg Thr Cys Lys Lys Glu Leu Thr Arg Lys Asp
 530 535 540
 60 Thr Ile Met Ala His Val Thr Glu Phe His Asn Gly His Arg Tyr Phe
 545 550 555 560
 65 Tyr Glu Met Asp Glu Val Glu Gly Glu Thr Leu Pro Ser Ser Ser Thr
 565 570 575
 Thr Leu Asp Asn Leu Thr Ala Asn Lys Pro Ser Ser Ala Ile Thr Val
 580 585 590
 Ile Asp His Ser Pro Ala Asn Ser Ser Pro Arg Gly Lys Trp Gln Cys
 595 600 605

ES 2 305 116 T3

5	Arg Ile Cys Glu Asp Met Phe Asp Ser Gln Glu Tyr Val Lys Gln His	610	615	620
10	Cys Met Ser Leu Ala Ser His Lys Phe His Arg Tyr Ser Cys Ala His	625	630	635 640
15	Cys Arg Lys Pro Phe His Lys Ile Glu Thr Leu Tyr Arg His Cys Gln	645	650	655
20	Asp Glu His Asp Asn Glu Ile Lys Ile Lys Tyr Phe Cys Gly Leu Cys	660	665	670
25	Asp Leu Ile Phe Asn Val Glu Glu Ala Phe Leu Ser His Tyr Glu Glu	675	680	685
30	His His Ser Ile Asp Tyr Val Phe Val Ser Glu Lys Thr Glu Thr Ser	690	695	700
35	Ile Lys Thr Glu Asp Asp Phe Pro Val Ile Glu Thr Ser Asn Gln Leu	705	710	715 720
40	Thr Cys Gly Cys Arg Glu Ser Tyr Ile Cys Lys Val Asn Arg Lys Glu	725	730	735
45	Asp Tyr Ser Arg Cys Leu Gln Ile Met Leu Asp Lys Gly Lys Leu Trp	740	745	750
50	Phe Arg Cys Ser Leu Cys Ser Ala Thr Ala Gln Asn Leu Thr Asp Met	755	760	765
55	Asn Thr His Ile His Gln Val His Lys Glu Lys Ser Asp Glu Glu Glu	770	775	780
60	Gln Gln Tyr Val Ile Lys Cys Gly Thr Cys Thr Lys Ala Phe His Asp	785	790	795 800
65	Pro Glu Ser Ala Gln Gln His Phe His Arg Lys His Cys Phe Leu Gln	805	810	815
	Lys Pro Ser Val Ala His Phe Gly Ser Glu Lys Ser Asn Leu Tyr Lys	820	825	830
	Phe Thr Ala Ser Ala Ser His Thr Glu Arg Lys Leu Lys Gln Ala Ile	835	840	845
	Asn Tyr Ser Lys Ser Leu Asp Met Glu Lys Gly Val Glu Asn Asp Leu	850	855	860

ES 2 305 116 T3

5 Ser Tyr Gln Asn Ile Glu Glu Glu Ile Val Glu Leu Pro Asp Leu Asp
 865 870 875 880
 Tyr Leu Arg Thr Met Thr His Ile Val Phe Val Asp Phe Asp Asn Trp
 885 890 895
 10 Ser Asn Phe Phe Gly His Leu Pro Gly His Leu Asn Gln Gly Thr Phe
 900 905 910
 15 Ile Trp Gly Phe Gln Gly Gly Asn Thr Asn Trp Lys Pro Pro Leu Asn
 915 920 925
 20 Cys Lys Ile Tyr Asn Tyr Leu Asn Arg Ile Gly Cys Phe Phe Leu His
 930 935 940
 25 Pro Arg Cys Ser Lys Arg Lys Asp Ala Ala Asp Phe Ala Ile Cys Met
 945 950 955 960
 His Ala Gly Arg Leu Asp Glu Gln Leu Pro Lys Gln Ile Pro Phe Thr
 965 970 975
 30 Ile Leu Ser Gly Asp Gln Gly Phe Leu Glu Leu Glu Asn Gln Phe Lys
 980 985 990
 35 Lys Thr Gln Arg Pro Ala His Ile Leu Asn Pro His His Leu Glu Gly
 995 1000 1005
 40 Asp Met Met Cys Ala Leu Leu Asn Ser Ile Ser Asp Thr Thr Lys Glu
 1010 1015 1020
 45 Cys Asp Ser Asp Asp Asn Met Gly Ala Lys Asn Thr Ser Ile Gly Glu
 1025 1030 1035 1040
 Glu Phe Ile Ser Thr Glu Asp Val Glu Leu Glu Glu Ala Ile Arg Arg
 1045 1050 1055
 50 Ser Leu Glu Gln Met
 1060
 55 <210> 12
 <211> 1061
 <212> PRT
 <213> *Homo sapiens*
 60 <400> 12
 65 Met Gly Asp Pro Gly Ser Glu Ile Ile Glu Ser Val Pro Pro Ala Gly

ES 2 305 116 T3

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

ES 2 305 116 T3

	260	265	270
5	Ser Gly Lys Asn His Phe His Gln Ser Phe Lys Leu Gly Asp Asn Lys 275	280	285
10	Gly Ile Ala His Pro Ile Ser Phe Pro Ser Phe Ala Lys Lys Leu Leu 290	295	300
15	Ile Ser Leu Cys Lys Asp Val Pro Phe Gln Val Lys Cys Val Ala Cys 305	310	315 320
	His Lys Thr Leu Arg Ser His Met Glu Leu Thr Ala His Phe Arg Val		
	325	330	335
20	His Cys Arg Asn Ala Gly Pro Val Ala Val Ala Glu Lys Ser Ile Thr 340	345	350
25	Gln Val Ala Glu Lys Phe Ile Leu Arg Gly Tyr Cys Pro Asp Cys Asn 355	360	365
30	Gln Val Phe Val Asp Glu Thr Ser Thr Gln Asn His Lys Gln Asn Ser 370	375	380
35	Gly His Lys Val Arg Val Ile Asn Ser Val Glu Glu Ser Val Leu Leu 385	390	395 400
40	Tyr Cys His Ser Ser Glu Gly Asn Lys Asp Pro Ser Ser Asp Leu His 405	410	415
45	Leu Leu Leu Asp Gln Ser Lys Phe Ser Ser Leu Lys Arg Thr Met Ser 420	425	430
50	Ile Lys Glu Ser Ser Ser Leu Glu Cys Ile Ala Ile Pro Lys Lys Lys 435	440	445
55	Met Asn Leu Lys Asp Lys Ser His Glu Gly Val Ala Cys Val Gln Lys 450	455	460
60	Glu Lys Ser Val Val Lys Thr Trp Phe Cys Glu Cys Asn Gln Arg Phe 465	470	475 480
65	Pro Ser Glu Asp Ala Val Glu Lys His Val Phe Ser Ala Asn Thr Met 485	490	495
	Gly Tyr Lys Cys Val Val Cys Gly Lys Val Cys Asp Asp Ser Gly Val 500	505	510
	Ile Arg Leu His Met Ser Arg Ile His Gly Gly Ala His Leu Asn Asn		

ES 2 305 116 T3

	515	520	525
5	Phe Leu Phe Trp Cys Arg Thr Cys Lys Lys Glu Leu Thr Arg Lys Asp 530	535	540
10	Thr Ile Met Ala His Val Thr Glu Phe His Asn Gly His Arg Tyr Phe 545	550	555 560
15	Tyr Glu Met Asp Glu Val Glu Gly Glu Thr Leu Pro Ser Ser Ser Thr 565	570	575
20	Thr Leu Asp Asn Leu Thr Ala Asn Lys Pro Ser Ser Ala Ile Thr Val 580	585	590
25	Ile Asp His Ser Pro Ala Asn Ser Ser Pro Arg Gly Lys Trp Gln Cys 595	600	605
30	Arg Ile Cys Glu Asp Met Phe Asp Ser Gln Glu Tyr Val Lys Gln His 610	615	620
35	Cys Met Ser Leu Ala Ser His Lys Phe His Arg Tyr Ser Cys Ala His 625	630	635 640
40	Cys Arg Lys Pro Phe His Lys Ile Glu Thr Leu Tyr Arg His Cys Gln 645	650	655
45	Asp Glu His Asp Asn Glu Ile Lys Ile Lys Tyr Phe Cys Gly Leu Cys 660	665	670
50	Asp Leu Ile Phe Asn Val Glu Glu Ala Phe Leu Ser His Tyr Glu Glu 675	680	685
55	His His Ser Ile Asp Tyr Val Phe Val Ser Glu Lys Thr Glu Thr Ser 690	695	700
60	Ile Lys Thr Glu Asp Asp Phe Pro Val Ile Glu Thr Ser Asn Gln Leu 705	710	715 720
65	Thr Cys Gly Cys Arg Glu Ser Tyr Ile Cys Lys Val Asn Arg Lys Glu 725	730	735
	Asp Tyr Ser Arg Cys Leu Gln Ile Met Leu Asp Lys Gly Lys Leu Trp 740	745	750
	Phe Arg Cys Ser Leu Cys Ser Ala Thr Ala Gln Asn Leu Thr Asp Met 755	760	765
	Asn Thr His Ile His Gln Val His Lys Glu Lys Ser Asp Glu Glu Glu		

ES 2 305 116 T3

	770		775		780	
5	Gln Gln Tyr Val Ile Lys Cys Gly Thr Cys Thr Lys Ala Phe His Asp					
	785		790		795	800
10	Pro Glu Ser Ala Gln Gln His Phe His Arg Lys His Cys Phe Leu Gln					
		805		810		815
15	Lys Pro Ser Val Ala His Phe Gly Ser Glu Lys Ser Asn Leu Tyr Lys					
		820		825		830
20	Phe Thr Ala Ser Ala Ser His Thr Glu Arg Lys Leu Lys Gln Ala Ile					
		835		840		845
25	Asn Tyr Ser Lys Ser Leu Asp Met Glu Lys Gly Val Glu Asn Asp Leu					
		850		855		860
30	Ser Tyr Gln Asn Ile Glu Glu Glu Ile Val Glu Leu Pro Asp Leu Asp					
		865		870		875
35	Tyr Leu Arg Thr Met Thr His Ile Val Phe Val Asp Phe Asp Asn Trp					
		885		890		895
40	Ser Asn Phe Phe Gly His Leu Pro Gly His Leu Asn Gln Gly Thr Phe					
		900		905		910
45	Ile Trp Gly Phe Gln Gly Gly Asn Thr Asn Trp Lys Pro Pro Leu Asn					
		915		920		925
50	Cys Lys Ile Tyr Asn Tyr Leu Asn Arg Ile Gly Cys Phe Phe Leu His					
		930		935		940
55	Pro Arg Cys Ser Lys Arg Lys Asp Ala Ala Asp Phe Ala Ile Cys Met					
		945		950		955
60	His Ala Gly Arg Leu Asp Glu Gln Leu Pro Lys Gln Ile Pro Phe Thr					
		965		970		975
65	Ile Leu Ser Gly Asp Gln Gly Phe Leu Glu Leu Glu Asn Gln Phe Lys					
		980		985		990
70	Lys Thr Gln Arg Pro Ala His Ile Leu Asn Pro His His Leu Glu Gly					
		995		1000		1005
75	Asp Met Met Cys Ala Leu Leu Asn Ser Ile Ser Asp Thr Thr Lys Glu					
		1010		1015		1020
80	Cys Asp Ser Asp Asp Asn Met Gly Ala Lys Asn Thr Ser Ile Gly Glu					

ES 2 305 116 T3

	1025	1030	1035	1040
5	Glu Phe Ile Ser Thr Glu Asp Val Glu Leu Glu Glu Ala Ile Arg Arg			
	1045 1050 1055			
10	Ser Leu Glu Glu Met			
	1060			
	<210> 13			
	<211> 19			
15	<212> ADN			
	<213> <i>Homo sapiens</i>			
	<400> 13			
20	atgggagacc cgggtcggg			19
	<210> 14			
	<211> 20			
25	<212> ADN			
	<213> <i>Homo sapiens</i>			
	<400> 14			
30	ttacatttcc tcaagacttc			20
	<210> 15			
35	<211> 30			
	<212> ADN			
	<213> Secuencia artificial			
40	<220>			
	<223> Descripción de secuencia artificial: cebador			
	<400> 15			
45	ctaggtaccg gacagcagga gcagtgggtgc			30
	<210> 16			
50	<211> 20			
	<212> ADN			
	<213> Secuencia artificial			
55	<220>			
	<223> Descripción de secuencia artificial: cebador			
	<400> 16			
60	ggcagtgagc tccatgtggg			20
65				