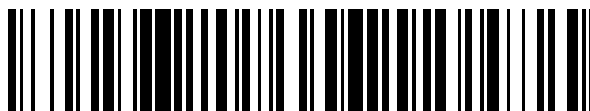


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 565 496**

51 Int. Cl.:

A61K 38/17 (2006.01)

A61K 38/39 (2006.01)

A61K 45/06 (2006.01)

A61K 48/00 (2006.01)

A61P 35/00 (2006.01)

A61K 31/513 (2006.01)

A61K 31/4745 (2006.01)

A61K 31/7048 (2006.01)

A61K 33/24 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.01.2004 E 12183770 (2)**

97 Fecha y número de publicación de la concesión europea: **13.01.2016 EP 2561887**

54 Título: **Sensibilizador a la terapia para el cáncer**

30 Prioridad:

14.01.2003 US 440009 P

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

05.04.2016

73 Titular/es:

**DANA-FARBER CANCER INSTITUTE (100.0%)
44 Binney Street
Boston, MA 02115, US**

72 Inventor/es:

**TAI, ISABELLE T. y
CHEN, LAN BO**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 565 496 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

Sensibilizador a la terapia para el cáncer

5 Campo de la invención

La invención se refiere a composiciones y métodos sensibilizantes a terapias para el cáncer.

Antecedentes

10

El motivo del fracaso de las terapias para el cáncer después de la inducción con quimioterapia o radioterapia sigue sin estar claro. Se han implicado muchos factores en la resistencia terapéutica, tales como la regulación positiva de las bombas efluentes de la p-glucoproteína de la familia de resistencia multifármaco (MDR) y se han descrito otras proteínas de MDR no clásicas (proteína asociada con la resistencia multifármaco, MRP; proteína de resistencia pulmonar, LRP) en una diversidad de cánceres (Lehnert, M. *Anticancer Res* 1998; 18:2225-2226; Ringborg, U. y Platz, A. *Acta Oncol* 1996; 5:76-80; Shea, T.C., Kelley, S.L., y Henner, W.D. *Cancer Res* 1998; 48:527-533). Desafortunadamente, muchos tumores que son intrínsecamente resistentes a la quimioterapia, tales como las neoplasias malignas gastrointestinales, tienen niveles relativamente bajos de los genes de MDR. Por ejemplo, solo el 23% de los tumores colorrectales primarios expresan MRP y el 65% expresan p-glucoproteína (Filipits M, Suchomel RW, Dekan G, Stiglbauer W, Haider K, Depisch D, Pirker R. *Br. J Cancer* 1997; 75: 208-212). Por lo tanto, la resistencia a agentes quimioterapéuticos no puede explicarse únicamente basándose en la activación y regulación positiva de genes de MDR conocidos. Los estudios también han demostrado que las mutaciones genéticas responsables de la tumorigénesis pueden contribuir también a la resistencia a fármacos. Por ejemplo, la pérdida de los genes de reparación de desemparejamiento de ADN (MMR) encontrados en el cáncer colorrectal hereditario no asociado a poliposis (HNPCC), se han asociado con una emergencia más rápida de resistencia clínica a fármacos (de las Alas, MM, S Aebi, D Fink, SB Howell, G Los. *J Natl Cane Inst* 1997; 89:1537-41; Lin X, Howell SB (1999). *Mol Pharmacol* 56:390-5). Las mutaciones en el gen de K-ras, detectadas en aproximadamente el 40% de los pólipos adenomatosos y adenocarcinomas, se asocian con una tasa aumentada de recaída, mortalidad y mala respuesta quimioterapéutica (Arber N, 1. Shapira, J. Ratan *et al. Gastroenterology* 2000; 118:1045-1050). Los genes implicados en la regulación del ciclo celular, tales como p21 y p27, han demostrado proteger a los tumores de sufrir apoptosis provocada por diversos agentes anticáncer (Waldman T, Lengauer C, Kinzler KW, Vogelstein B. *Nature* 1996, 381:713-716; St. Croix B, Florenes VA, Rak JW, Flanagan M, Bhattacharya N, Slingerland JM, Kerbel RS. *Nature Med* 1996, 2:1204-1210). Además, las moléculas de adhesión celular, tales como E-cadherina, confieren resistencia a las células cuando se exponen a agentes quimioterapéuticos (Skoudy A, Llosas MD, Garcia de Herreros A. *Biochem J* 1996).

40

Los mecanismos implicados en la resistencia terapéutica, por lo tanto, parecen ser muy complejos. Pruebas recientes sugieren que la selectividad de la quimioterapia para los relativamente pocos tumores que llegan a curarse mediante fármacos depende, en gran medida, de su fácil susceptibilidad para sufrir apoptosis, es decir, para eliminarse a sí mismos (Making G, *Expert Opin Ther Targets*. 2002 6(1):73-84; Johnstone RW, Ruefli AA, Lowe SW, *Cell*. 2002 108(2):153-64; Kamesaki H, *int J Hematol*. 1998 68(1):29-43).

45

La proteína ácida secretada y rica en cisteína (SPARC) pertenece a una familia de proteínas extracelulares, denominadas proteínas matricelulares. Desde su identificación y clonación, el papel funcional de SPARC sigue sin estar claro. Su alta conservación evolutiva sugiere un papel fisiológico importante para esta proteína (Iruela-Arispe ML, Lane TF, Redmond D, Reilly M, Bolender RP, Kavanagh TJ, Sage EH. *Mol Biol Cell*. marzo de 1995; 6(3):327-43). Los estudios iniciales mostraron que SPARC es importante para la mineralización ósea (Termine JD, Kleinman HK, Whitson SW, Conn KM, McGarvey ML, Martin GR. *Cell*. octubre de 1981; 26(1 Pt 1):99-105). Aunque SPARC se expresa a altos niveles en el tejido óseo, también se distribuye ampliamente en otros tejidos y tipos celulares (Maillard, C., *et al.*, *Bone*, 13:257-264 (1992)). Su papel se ha extendido para incluir la remodelación ósea (Latvala T, Puolakkainen P, Vesaluoma M, Tervo T. *Exp Eye Res*. noviembre de 1996;63(5):579-84; Kehn RJ Jr, Swords NA, Orfeo T, Mann KG. *J Biol Chem*. 2 de diciembre de 1994;269(48):30147-53); la migración celular endotelial (Hasselaar P, Sage EH. *J Cell Biochem*. jul 1992;49(3):272-83), la morfogénesis (Mason IJ, Murphy D, Munke M, Francke U, Elliott RW, Hogan BL. *EMBO J*. agosto de 1986; 5(8):18831-7; Strandjord TP, Sage EH, Clark JG. *Am J Respir Cell Mol Biol*. septiembre de 1995; 13(3):279-87), y la angiogénesis (Kupprion C, Motamed K, Sage EH. *J Biol Chem*. 6 de nov de 1998; 273(45):29635-40; Lane TF, Iruela-Arispe ML, Johnson RS, Sage EH. *J Cell Biol*. mayo de 1994; 125(4):929-43). También se ha demostrado que SPARC tiene un efecto antiproliferativo en las células endoteliales, células mesangiales, fibroblastos y células musculares lisas (Sage EH. *Biochem Cell Biol*.jul 1992; 70(7):579-92).

60

También se ha identificado a SPARC en tumores en experimentos *in vitro* (Schulz, A., *et al.*, *Am. J. Pathol.*, 132:233-238 (1988); Porter, P.L., *et al.*, *J. Histochem. Cytochem.*, 43:791- 800 (1995)). Hay pruebas opuestas de que SPARC puede funcionar bien como un oncogén, tal como se sugiere por estudios en melanoma (Ledda MF, Adris S, Bravo AT, Kairiyama C, Bover L, Chemajovsky Y, Mordoh J, Podhajcer OL. *Nat Med*. febrero de 1997; 3(2):171-6) o como un supresor tumoral, tal como se demuestra por su fuerte inhibición del crecimiento en fibroblastos embrionarios de pollo transformados con vJun-ml y v-Src (Vial E, Castellazzi M. *Oncogene*. 30 de marzo de 2000; 19(14):1772-82).

65

Aunque las propiedades inhibitorias del crecimiento de SPARC se han demostrado principalmente en células primarias, tales como endoteliales, fibroblasto, mesangiales y células musculares lisas, esto puede contribuir también al papel de SPARC en la tumorigénesis. También se ha demostrado que SPARC tiene propiedades invasivas de tumores. Se ha observado expresión variable de SPARC en una diversidad de cánceres. Se han detectado niveles de expresión mayores en el cáncer de mama (Bellahcene A, Castronovo V. *Am J Pathol.* enero de 1995; 146(1):95-100), cáncer de esófago (Porte H, Triboulet JP, Kotelevets L, Carrat F, Prevot S, Nordlinger B, DiGioia Y, Wurtz A, Comoglio P, Gespach C, Chastre E. *Clin Cancer Res.* junio de 1998; 4(6):1375-82), carcinoma hepatocelular (Le Bail B, Faouzi S, Boussarie L, Guirouilh J, Blanc JF, Carles J, Bioulac-Sage P, Balabaud C, Rosenbaum J. *J Pathol.* septiembre de 1999; 189(1):46-52), y de próstata (Thomas R, True LD, Bassuk JA, Lange PH, Vessella RL. *Clin Cancer Res* 2000; 6:1140- 1149). Sin embargo, se han observado resultados contradictorios con cánceres de ovario (Brown TJ, Shaw PA, Karp X, Huynh MH, Begley, Ringuette MJ. *Gynecol Oncol* 1999; 75: 25-33; Paley PJ, GoffBA, Gown AM, Greer BE, Sage EH. *Gynecol Oncol* 2000; 78: 336-341; Yiu GK, Chan WY, Ng SW, Chan PS, Cheung KK, Berkowitz RS, Mok SC. *Am J Pathol* 2001; 159:609-622), y cánceres colorrectales (Porte H, Chastre E, Prevot S, Nordlinger B, Empereur S, Basset P, Chambon P, Gespach C. *Int J Cancer* 1995; 64: 70-75; Lussier C, Sodek J, Beaulieu JF. *J Cell Biochem.* 2001;81(3):463-76).

Recientemente, se ha sugerido que SPARC está implicada en la inducción de la apoptosis de las células de cáncer de ovario (Yiu GK, Chan WY, Ng SW, Chan PS, Cheung KK, Berkowitz RS, Mok SC. *Am J Pathol* 2001; 159:609-622). Yiu *et al.* (2001, anteriormente citado) han demostrado había regulación negativa de la expresión de SPARC después de la transformación maligna, y que había propiedades antiproliferativas de SPARC en células ováricas tanto normales como cancerosas. Yiu *et al.* (2001, anteriormente citado) proporcionaron pruebas adicionales de que la exposición exógena solo a SPARC fue capaz de inducir la apoptosis en células de cáncer de ovario. Sin embargo, no se ha demostrado que los especímenes de patología humanos con altos niveles de expresión de SPARC tengan un mayor número de células apoptóticas.

Las estrategias de quimiosensibilización usan la administración de un fármaco o agente para hacer que las células cancerosas sean más susceptibles a un segundo fármaco. La quimiosensibilización puede permitir una dosis menor, menos tóxica o restaurar la actividad de fármacos a los que las células cancerosas se han hecho resistentes. Una serie de estrategias que se han usado para quimiosensibilizar. Por ejemplo, Stein *et al.*, demostraron que la expresión endógena de genes de citocinas en células tumorales, tales como TNF e IL-2, pueden modular la resistencia a múltiples fármacos *in vitro*. Esta modulación potencia la susceptibilidad de las células a los fármacos citotóxicos. (*Journal of the National Cancer Institute*, 1996; 88 (19) 1383-92).

El documento WO0202771 divulga un nuevo polipéptido de hSPARC-h1 y sus aplicaciones potenciales en la remodelación de tejidos, reparación de tejidos y modulación general de diversas actividades de factor de crecimiento.

La Patente de Estados Unidos n.º 6.387.664 proporciona una proteína de fusión de SPARC que puede obtenerse fusionando SPARC a tioredoxina, que puede usarse para investigación básica en neurobiología y/o para tratar diversas neuropatologías.

La Patente de Estados Unidos n.º 6.239.326 proporciona un modelo de ratón transgénico deficiente para SPARC para ensayar fármacos que promuevan o retrasen la curación de heridas y que traten o prevengan las cataratas, la diabetes mellitus o la osteoporosis.

Sumario de la invención

La materia-objeto que no esté abarcada por el alcance de las reivindicaciones no forma parte de la invención reivindicada en el presente documento. La invención se basa en el descubrimiento de que SPARC sensibiliza a la terapia para el cáncer.

La presente invención proporciona el uso de un polinucleótido de la familia de SPARC que codifica un polipéptido de la familia de SPARC con la secuencia de SEC ID N°: 1 en la preparación de un medicamento para aumentar la sensibilidad o reducir la resistencia de un mamífero al que se le ha diagnosticado cáncer a un tratamiento terapéutico para el cáncer en el que dicho mamífero muestra resistencia a dicho tratamiento terapéutico.

En el presente documento se divulga una composición que comprende un polipéptido de la familia de SPARC y un agente quimioterapéutico.

Se enseña en el presente documento una célula resistente a la quimioterapia que comprende un polinucleótido de la familia de SPARC recombinante.

También se enseña en el presente documento una célula recombinante que comprende una región de control de la transcripción heteróloga asociada operativamente con un polinucleótido de la familia de SPARC.

En otro aspecto, la presente invención proporciona un polinucleótido de la familia de SPARC que codifica un polipéptido de la familia de SPARC con la secuencia de SEC ID N°: 1 para su uso en el aumento de la sensibilidad o reducción de la resistencia de un mamífero al que se le ha diagnosticado cáncer a un tratamiento terapéutico para el cáncer, en el que dicho mamífero muestra resistencia a dicho tratamiento terapéutico.

También se enseña en el presente documento un método *ex vivo* para sensibilizar a un mamífero al que se le ha diagnosticado cáncer a un tratamiento terapéutico, comprendiendo el método: administrar a un mamífero una cantidad eficaz de una célula que comprenda un polipéptido de la familia de SPARC o de un polinucleótido que codifica un polipéptido de la familia de SPARC; en el que la célula produce una cantidad aumentada del polipéptido SPARC.

En el presente documento se enseña un método para sensibilizar *ex vivo* a un mamífero al que se le ha diagnosticado cáncer a un tratamiento terapéutico, comprendiendo el método: (1) obtener una muestra de cáncer del mamífero; (2) poner en contacto la muestra de sangre con una cantidad eficaz de una composición que comprende un polipéptido de la familia de SPARC o de un polinucleótido que codifica un polipéptido de SPARC; y (3) devolver la muestra de cáncer después de la puesta en contacto de (2) al mamífero.

También se enseña en el presente documento un método para sensibilizar a una muestra de cáncer a un tratamiento terapéutico, comprendiendo el método poner en contacto la muestra de cáncer con una cantidad eficaz de una composición que comprende un polipéptido de la familia de SPARC o un polinucleótido que codifica un polipéptido de SPARC.

Tal como se divulga en el presente documento, la muestra de cáncer puede ser una muestra de células o de tejido.

Tal como se divulga en el presente documento, la muestra de cáncer puede transfectarse o infectarse con el polinucleótido de SPARC.

En el presente documento se enseña un método para evaluar una primera célula de cáncer respecto de su resistencia a un tratamiento terapéutico que comprende: (a) medir el nivel de expresión de un ARNm o polipéptido de la familia de SPARC, o el nivel extracelular de un polipéptido de la familia de SPARC en la primera célula cancerosa; y (b) comparar el nivel de expresión o el nivel extracelular obtenido en (a) con el nivel de expresión del ARNm o el polipéptido de la familia de SPARC, o el nivel extracelular del polipéptido de la familia de SPARC en una segunda célula cancerosa que no muestra resistencia al tratamiento terapéutico; en el que un nivel menor de expresión o nivel extracelular en (a) es indicativo de que la primera célula cancerosa es resistente al tratamiento terapéutico.

También se enseña en el presente documento un método para evaluar una primera célula cancerosa respecto de su resistencia a un tratamiento terapéutico, que comprende: (a) medir el nivel de expresión de un ARNm o polipéptido de la familia de SPARC, o el nivel extracelular de un polipéptido de la familia de SPARC en la primera muestra de cáncer; (b) medir el nivel de expresión del ARNm o polipéptido de la familia de SPARC en una segunda muestra de cáncer que no muestra resistencia al tratamiento terapéutico; (c) comparar los niveles de expresión o los niveles extracelulares obtenidos en (a) y (b), donde un nivel de expresión o nivel extracelular menor en (a) es indicativo de que la primera muestra de cáncer es resistente al tratamiento terapéutico.

Tal como se divulga en el presente documento, la primera muestra puede ser de un primer mamífero y el nivel de expresión o nivel extracelular menor en (a) es indicativo además de que el primer mamífero es resistente al tratamiento terapéutico.

Tal como se divulga en el presente documento, la segunda muestra de cáncer puede ser de un primer mamífero que proporciona la primera muestra de cáncer.

De manera adecuada, la segunda muestra de cáncer puede ser de un segundo mamífero que sea diferente al primer mamífero que proporciona la primera muestra de cáncer.

El primer mamífero y el segundo mamífero pueden haber sido diagnosticados del mismo cáncer.

El nivel de expresión del ARNm de la familia de SPARC puede medirse mediante reacción en cadena de la polimerasa, micromatriz de ADN o transferencia de Northern.

Tal como se divulga en el presente documento, el nivel de expresión o extracelular del polipéptido de la familia de SPARC puede medirse mediante inmunotransferencia o ensayo inmunoabsorbente ligado a enzimas (ELISA).

En el presente documento se divulga un método para identificar a un agente que modula la expresión de un ARNm o polipéptido de la familia de SPARC, o la secreción de un polipéptido de la familia de SPARC, que comprende: (a) medir el nivel de expresión del ARNm o polipéptido de la familia de SPARC, o el nivel extracelular del polipéptido de la familia de SPARC; (b) poner en contacto un agente candidato con la muestra; (c) después de la puesta en

contacto de (b), medir la expresión o el nivel extracelular del polipéptido de la familia de SPARC en la muestra de (b); (d) comparar los niveles de expresión o los niveles extracelulares en (a) y (c), donde un nivel de expresión o nivel extracelular diferencial en (a) y (c) indica que el agente candidato es un agente que modula la expresión del ARNm o polipéptido de la familia de SPARC, o la secreción del polipéptido de la familia de SPARC.

En el presente documento se divulga un método para identificar un agente que sensibiliza a una muestra de cáncer a un tratamiento terapéutico, que comprende: (a) medir el nivel de expresión de un ARNm o polipéptido de la familia de SPARC, o el nivel extracelular de un polipéptido de la familia de SPARC en la muestra de cáncer; (b) poner en contacto un agente candidato con la muestra de cáncer; (c) después de la puesta en contacto de (b), medir el nivel de expresión del ARNm o polipéptido de la familia de SPARC, o el nivel extracelular del polipéptido de la familia de SPARC en la muestra de cáncer de (b); (d) comparar los niveles de expresión o los niveles extracelulares obtenidos en (a) y (c), donde un nivel de expresión o un nivel extracelular en (c) indica que el agente candidato es un agente que sensibiliza a una muestra de cáncer a un tratamiento terapéutico.

Tal como se divulga en el presente documento, la muestra de cáncer del método objeto puede ser de un mamífero al que se le ha diagnosticado cáncer, y el nivel de expresión aumentado o nivel extracelular es indicativo además de que el agente candidato es un agente que sensibiliza al mamífero a un tratamiento terapéutico.

En el presente documento se divulga un método para determinar un protocolo de tratamiento terapéutico para un primer mamífero al que se le ha diagnosticado cáncer, que comprende: (a) determinar si la expresión de un ARNm o polipéptido de la familia de SPARC o el nivel extracelular de un polipéptido de la familia de SPARC es menor en una primera muestra del primer mamífero que en una segunda muestra que no muestre resistencia al tratamiento terapéutico; y (b) si (a) es positivo, aumentar la fuerza del tratamiento terapéutico para el primer mamífero para aumentar la respuesta al tratamiento.

Preferentemente, el polinucleótido de SPARC de la composición y método objeto es un vector de expresión.

Más preferentemente, el vector de expresión es un plásmido o un vector viral.

Tal como se enseña en el presente documento, el vector plasmídico es pcDNA3.1.

El polipéptido de la familia de SPARC o un polinucleótido que codifica un polipéptido de la familia de SPARC tal como se enseña en el presente documento incluye: (a) un polipéptido de SPARC que se selecciona entre el grupo que consiste en las SEC ID N°: 1-17; (b) un polipéptido que tiene una secuencia de aminoácidos con al menos un 60% de homología con el polipéptido de la familia de SPARC de (a); (c) un fragmento polipeptídico de (a)-(b) donde el fragmento tiene al menos 50 aminoácidos de longitud; (d) un polipéptido de fusión que comprende el polipéptido de (a), (b), o (c); (e) un polinucleótido que codifica al polipéptido de (a), (b), (c) o el polipéptido de fusión de (d); o (f) un polinucleótido que hibrida con el polinucleótido de (e) en condiciones de hibridación rigurosas.

Preferentemente, el polipéptido o polinucleótido de la familia de SPARC se selecciona entre: SMOC-1, SPARC, hevina, SCI, QR-1, proteínas similares a folistatina (TSC-36), testican.

Preferentemente, el agente terapéutico de la composición y método objeto es un agente quimioterapéutico o un agente de radioterapia.

Más preferentemente, el agente terapéutico se selecciona entre el grupo que consiste en los agentes listados en la tabla 1.

En una realización, el mamífero del uso objeto muestra resistencia al tratamiento terapéutico.

En una realización, el tratamiento terapéutico es quimioterapia o radioterapia.

Tal como se enseña en el presente documento, la muestra de cáncer del método objeto puede ser una célula o una muestra de tejido.

Preferentemente, la célula o la muestra de tejido se lisan antes de medir la expresión o nivel extracelular de un polipéptido o polinucleótido de la familia de SPARC (por ejemplo, ARNm).

Más preferentemente, se obtiene un polinucleótido o extracto de polipéptido de la célula o muestra de tejido antes de medir la expresión o nivel extracelular de un polipéptido o polinucleótido de la familia de SPARC (por ejemplo, ARNm).

Breve descripción de los dibujos

Las características anteriores de la invención se entenderán más fácilmente por referencia a la siguiente descripción detallada, tomada con referencia a los dibujos adjuntos:

La figura 1 es una estructura modular de SPARC humano y la localización y funciones de péptidos sintéticos. Se muestran tres dominios y sus números de resto tal como se describe en Yan y Sage, 1999, J. Histochem. & Cytochem., 47(12):1495-1505.

5 La figura 2 es una organización de dominios de diversas proteínas de la familia de SPARC. FS representa el dominio similar a folistatina, TY el dominio similar a tiroglobulina, EC el dominio de unión a calcio extracelular, tal como se describe en Vannahme *et al.*, 2002, J. Biol. Chem., 277(41):37977-37986. Los dominios sin homología con otras proteínas se muestran como cajas abiertas. No se muestran los péptidos de señal.

10 La figura 3 es una fotografía que muestra los ensayos de formación de colonias de células sensibles y resistentes a la quimioterapia de acuerdo con una realización de la invención.

La figura 4 es una fotografía que muestra los ensayos TUNEL de células sensibles y resistentes a la quimioterapia de acuerdo con una realización de la invención.

15 La figura 5 (A y B) es una fotografía que muestra el nivel disminuido de polipéptido de SPARC en líneas celulares resistentes a la quimioterapia, de acuerdo con una realización de la invención.

20 La figura 6 es un ensayo TUNEL que muestra la respuesta de las células MIP101 resistentes a SPARC exógeno para revertir el fenotipo resistente, de acuerdo con una realización de la invención.

La figura 7 es un ensayo de inmunotransferencia que muestra células recombinantes que expresan polipéptido de SPARC, de acuerdo con una realización de la invención.

25 La figura 8 es un análisis FACS que muestra poblaciones celulares inducidas para sufrir apoptosis después de la exposición a agentes quimioterapéuticos, de acuerdo con una realización de la invención.

La figura 9 es una gráfica que muestra el porcentaje de apoptosis de células después de la exposición a agentes quimioterapéuticos, de acuerdo con una realización de la invención.

30 La figura 10 es una gráfica que muestra la respuesta de transfectantes de SPARC a agentes quimioterapéuticos, de acuerdo con una realización de la invención.

35 La figura 11 es una presentación gráfica que muestra la regresión tumoral completa en animales a los que se han trasplantado transfectantes de SPARC después de 6 ciclos de quimioterapia en dos animales, de acuerdo con una realización de la invención.

40 La figura 12 es una figura que contiene las secuencias polinucleotídicas y polipeptídicas de los miembros de la familia de SPARC, de acuerdo con una realización de la invención.

La figura 13 es una figura que muestra el alineamiento de secuencias entre diferentes polipéptidos y polinucleótidos de la familia de SPARC, de acuerdo con una realización de la invención.

45 La figura 14 muestra el efecto de la quimioterapia en xenoinjertos de tumor de células que sobreexpresan SPARC, de acuerdo con una realización de la invención.

La figura 15 muestra el efecto de la radioterapia en xenoinjertos de tumor de células que sobreexpresan SPARC, de acuerdo con una realización de la invención.

50 La figura 16 muestra el tratamiento de xenoinjertos de tumor de MIP101 con terapia combinada con SPARC por vía intraperitoneal, de acuerdo con una realización de la invención.

La figura 17 muestra el tratamiento de xenoinjertos de tumor de MIP101 con terapia combinada con SPARC por vía subcutánea, de acuerdo con una realización de la invención.

55 La figura 18 muestra el tratamiento de xenoinjertos de tumor de MIP/5FU con terapia combinada con SPARC, de acuerdo con una realización de la invención.

60 La figura 19 muestra los niveles de ARNm y proteína de SPARC humano en líneas celulares de cáncer colorrectal sensibles y resistentes a la quimioterapia, de acuerdo con una realización de la invención.

La figura 20 muestra la expresión de proteína de SPARC en epitelio colónico humano, de acuerdo con una realización de la invención.

65 La figura 21 muestra la evaluación del efecto de SPARC para influenciar la sensibilidad de las células a la quimioterapia, de acuerdo con una realización de la invención.

Descripción detallada

La invención se basa en la familia de SPARC y en la sensibilización a la terapia para el cáncer.

5 Definiciones

Tal como se usa en el presente documento, un "polipéptido de la familia de SPARC" se refiere a un polipéptido (incluyendo un fragmento o variante del mismo) de una familia de proteínas extracelulares. Esta familia de proteínas extracelulares incluye SPARC y otros miembros de la familia, tales como SMOC-1, hevina, SCI, QR-1, proteínas similares a folistatina (TSC-36) y testicano (véase, por ejemplo, Vannahme *et al.*, (2002), J. Biol. Chem., 277(41): 37977-37986; Johnston, I. G., Paladino, T., Gurd, J. W., y Brown, I. R. (1990) Neuron 2, 165-176; Guemah, M., Crisanti, P., Laugier, D., Dezelee, P., Bidou, L., Pessac, B., y Calothy, G. (1991) Proc. Natl. Acad. Sci. U.S. A. 88, 4503-4507; Shibamura, M., Mashimo, J., Mita, A., Kuroki, T., y Nose, K. (1993) Eur. J. Biochem. 217, 13-19; Alliel, P.M., Perin, J. P., Jolles, P., y Bonnet, F. J. (1993) Eur. J. Biochem. 214, 347-350. Un polipéptido de la familia de SPARC está compuesto típicamente de tres dominios plegados de manera independiente (Yan y Sage, 1999, J. Histochem. & Cytochem., 47(12):1495-1505). El dominio N-terminal (por ejemplo, dos Glu₃ y Glu₄ N-terminales en SPARC), está cargado negativamente, el segundo dominio (por ejemplo, los restos 53-137 en SPARC) es homólogo a folistatina (FS)1 con 10 cisteínas en un patrón típico, el dominio extracelular de unión a calcio (EC) (por ejemplo, restos 138-286 en SPARC) tiene dos motivos de unión a calcio de mano de EF, cada uno con un ión de calcio unido en la estructura de rayos X (Maurer, P., Hohenadl, C., Hohenester, E., Gohring, W., Timpl, R., y Engel, J. (1995) J. Mol. Biol. 253, 347-357; Hohenester, E., Maurer, P., Hohenadl, C., Timpl, R., Jansonius, J. N., y Engel, J. (1996) Nat. Struct. Biol. 3, 67-73).

La expresión "polipéptido de la familia de SPARC", tal como se enseña en el presente documento, incluye el polipéptido de longitud completa o un fragmento del mismo, un polipéptido de tipo silvestre o cualquier variante del mismo. Un "polinucleótido de la familia de SPARC" es una molécula de polinucleótido (por ejemplo, ADN o ARNm) que codifica un polipéptido de SPARC, o un fragmento del mismo. (a) un polipéptido o gen de la familia de SPARC seleccionado entre el grupo que consiste en las SEC ID N°: 1-17; (b) un polipéptido que tiene una secuencia de aminoácidos al menos un 60% homóloga al polipéptido de la familia de SPARC en (a) o un gen que codifica al polipéptido de al menos un 60% de homología; (c) un fragmento polipeptídico de (a)-(b) en el que el fragmento tiene al menos 50 aminoácidos de longitud; (d) un polipéptido de fusión que comprende el polipéptido de (a), (b), o (c); (e) un polinucleótido que codifica al polipéptido de (a), (b), (c) o el polipéptido de fusión de (d); o (f) un polinucleótido que hibrida con el polinucleótido de (e) en condiciones de hibridación rigurosas.

La expresión "polipéptido de SPARC" se refiere a las SEC ID N°: 1-17, y la expresión "gen de SPARC" a las SEC ID N°: 18-34, las secuencias de nucleótido correspondientes a las SEC ID N°: 1-17. Se contempla que las variaciones de estas secuencias que retengan la actividad biológica de SPARC sean equivalentes de estas secuencias. La actividad biológica del gen de SPARC es que se regula negativamente en células resistentes a la quimioterapia. El gen también puede sobreexpresarse en células que sean sensibles a la quimioterapia. La actividad biológica del polipéptido de SPARC es que en células resistentes a la quimioterapia, las sensibiliza a la quimioterapia.

Con respecto a un polipéptido miembro de la familia de SPARC, se contempla que las variaciones de sus secuencias que retengan la actividad biológica del miembro de la familia sean equivalentes de estas secuencias. La actividad biológica de un miembro de la familia del gen de SPARC es que el gen se regula negativamente en células resistentes a la quimioterapia, es decir, su expresión disminuye en al menos un 25%, por ejemplo, 30%, 40%, 50%, 75%, 100% (1 vez), 2 veces, 4 veces, o 5 veces o más, en comparación con células resistentes a la quimioterapia. El gen también puede sobreexpresarse en células que sean sensibles a la quimioterapia. La actividad biológica del polipéptido miembro de la familia de SPARC es que en células resistentes a la quimioterapia, las sensibiliza a la quimioterapia, es decir, aumenta la respuesta a la quimioterapia en al menos un 25%, por ejemplo, 30%, 40%, 50%, 60%, 70%, 80%, o más, a 2 veces, 3 veces, 4 veces, 5 veces, 10 veces, 15 veces, 20 veces o más, en comparación con la sensibilidad al tratamiento en ausencia del polipéptido de la familia de SPARC.

Tal como se define en el presente documento, un "tejido" es un agregado de células que efectúa una función concreta en un organismo. El término "tejido", tal como se usa en el presente documento, se refiere a material celular de una región fisiológica concreta. Las células en un tejido particular pueden comprender varios tipos celulares diferentes. Un ejemplo no limitante de esto pueden ser células cerebrales, que además comprenden neuronas y células gliales, así como células endoteliales capilares y células sanguíneas.

La expresión "tipo celular" o "tipo tisular" se refiere al tejido de origen, por ejemplo, a partir del cual se desarrolla un tumor. Dichos tejidos (tipos celulares) incluyen, por ejemplo, sin limitación, sangre, colorrectal, de mama, esofágico, hepatocelular, de próstata, ovárico, tiroideos, páncreas, de útero, testículos, pituitaria, riñón, de estómago, esófago y recto.

Tal como se usa en el presente documento, el término "cáncer" se refiere a una enfermedad o trastorno proliferativo causada o caracterizada por la proliferación de células que han perdido la susceptibilidad para el control del crecimiento normal. El término "cáncer", tal como se usa en la presente solicitud, incluye tumores y cualquier otro

tipo de trastorno proliferativo. Los cánceres del mismo tipo tisular se originan en el mismo tejido, y pueden dividirse en diferentes subtipos basándose en sus características biológicas. El cáncer puede seleccionarse entre uno o más del grupo que consiste en: melanoma, leucemia, astrocitoma, glioblastoma, linfoma, glioma, linfoma de Hodgkin, leucemia linfocítica crónica y cáncer de páncreas, de mama, tiroides, de ovario, de útero, testículos, pituitaria, riñón, de estómago, esófago y recto.

Tal como se usa en el presente documento, el término "sensibilizar" se refiere a una sensibilidad aumentada o reducción de la resistencia de una muestra de cáncer o de un mamífero que responde a un tratamiento terapéutico. Una sensibilidad aumentada o una sensibilidad reducida a un tratamiento terapéutico se mide de acuerdo con un método conocido en la técnica para el tratamiento particular y con métodos descritos más adelante en el presente documento, incluyendo, pero sin limitación, ensayos de proliferación celular (Tanigawa N, Kern DH, Kikasa Y, Morton DL, Cancer Res 1982; 42: 2159-2164), ensayos de muerte celular (Weisenthal LM, Shoemaker RH, Marsden JA, Dill PL, Baker JA, Moran EM, Cancer Res 1984; 94:161-173; Weisenthal LM, Lippman ME, Cancer Treat Rep 1985; 69: 615-632; Weisenthal LM, En: Kaspers GJL, Pieters R, Twentymann PR, Weisenthal LM, Veerman AJP, eds. Drug Resistance in Leukemia and Lymphoma. Langhorne, PA: Harwood Academic Publishers, 1993: 415-432; Weisenthal LM, Contrib Gynecol Obstet 1994; 19: 82-90). La sensibilidad o resistencia también puede medirse en animales midiendo la reducción del tamaño tumoral durante un periodo de tiempo, por ejemplo, 6 meses para seres humanos y 4-6 semanas para ratones. Una composición o un método sensibiliza la respuesta a un tratamiento terapéutico si el aumento en la sensibilidad del tratamiento o la reducción en la resistencia es del 25% o más, por ejemplo, 30%, 40%, 50%, 60%, 70%, 80%, o más, a 2 veces, 3 veces, 4 veces, 5 veces, 10 veces, 15 veces, 20 veces o más, en comparación con la sensibilidad o resistencia al tratamiento en ausencia de dicha composición o método. La determinación de la sensibilidad o resistencia a un tratamiento terapéutico es rutinaria en la técnica y se encuentra dentro de las capacidades de un profesional sanitario experto habitual.

Tal como se usa en el presente documento, el término "administrar" o "administración" se refiere a introducir por cualquier medio una composición (por ejemplo, un agente terapéutico) en el cuerpo de un mamífero para prevenir o tratar una enfermedad o afección (por ejemplo, cáncer).

Tal como se usa en el presente documento, los términos "tratar", "tratamiento", "terapia" y "tratamiento terapéutico", tal como se usan en el presente documento se refieren a terapia curativa, terapia profiláctica o terapia preventiva. Un ejemplo de "terapia preventiva" es la prevención o alivio de una enfermedad diana (por ejemplo, cáncer) o una afección relacionada con la misma. Aquellos que necesitan tratamiento incluyen aquellos que ya tengan la enfermedad o afección así como aquellos propensos a tener la enfermedad o afección que se va a prevenir. Los términos "tratando", "tratamiento", "terapia" y "tratamiento terapéutico", tal como se usa en el presente documento también describen la gestión y cuidado de un mamífero con el fin de combatir una enfermedad, o afección relacionada, e incluye la administración de una composición para aliviar los síntomas, efectos secundarios, u otras complicaciones de la enfermedad o afección. El tratamiento terapéutico para el cáncer incluye, pero sin limitación, cirugía, quimioterapia, radioterapia, terapia génica, e inmunoterapia.

Por "cantidad terapéuticamente eficaz" se entiende una cantidad que alivia (hasta cierto punto, a juicio de un médico experto) uno o más síntomas de la enfermedad o afección en un mamífero. Además, por "cantidad terapéuticamente eficaz" se entiende una actividad que devuelva a la normalidad, ya sea parcial o completamente, parámetros fisiológicos y bioquímicos asociados con o causantes de una enfermedad o afección. Un profesional sanitario experto en la materia puede determinar la cantidad terapéuticamente eficaz de una composición para tratar o prevenir una patología concreta o trastorno cuando se administra, tal como por vía intravenosa, subcutánea, intraperitoneal, por vía oral, o mediante inhalación. La cantidad precisa de la composición necesaria para que sea terapéuticamente eficaz dependerá de numerosos factores, por ejemplo, la actividad específica del agente activo, el dispositivo de administración empleado, las características físicas del agente, el fin de la administración, además de muchas consideraciones específicas del paciente. La determinación de la cantidad de una composición que ha de administrarse para que sea terapéuticamente eficaz es rutinaria en la técnica y se encuentra dentro de las capacidades de un profesional sanitario experto habitual.

Tal como se usa en el presente documento, el término "agente" o "fármaco" o "agente terapéutico" se refiere a un compuesto químico, a una mezcla de compuestos químicos, a una macromolécula biológica, o a un extracto producido a partir de materiales biológicos, tales como bacterias, plantas, hongos, o células animales (en particular de mamífero) o tejidos que se sospecha que tengan propiedades biológicas. El agente o fármaco puede purificarse, purificarse sustancialmente o purificarse parcialmente. Un "agente", de acuerdo con la presente invención, también incluye un agente de radioterapia.

Tal como se usa en el presente documento, "modulación" o "modular" significa que una respuesta deseada/seleccionada es más eficaz (por ejemplo, al menos un 10%, 20%, 40%, 60% o más), más rápida (por ejemplo, al menos un 10%, 20%, 40%, 60% o más), de mayor magnitud (por ejemplo, al menos un 10%, 20%, 40%, 60% o mayor), y/o inducida más fácilmente (por ejemplo, al menos un 10%, 20%, 40%, 60% o más) en presencia de un agente que en ausencia del agente.

Tal como se usa en el presente documento, la expresión "resistencia a un tratamiento terapéutico" se refiere a una resistencia adquirida o natural de una muestra de cáncer o de un mamífero a una terapia, es decir, que no responde a o que tiene una respuesta reducida o limitada al tratamiento terapéutico, por ejemplo, que tiene una respuesta reducida a un tratamiento terapéutico en un 25% o más, por ejemplo, 30%, 40%, 50%, 60%, 70%, 80%, o más, a 2 veces, 3 veces, 4 veces, 5 veces, 10 veces, 15 veces, 20 veces o más. La reducción en la respuesta se mide comparando con la misma muestra de cáncer o mamífero antes de que se adquiriera la resistencia, o comparando con una muestra de cáncer diferente o mamífero que se sabe que no tiene resistencia al tratamiento terapéutico. Una resistencia adquirida típica a la quimioterapia se denomina "resistencia a múltiples fármacos". La resistencia a múltiples fármacos puede estar mediada por P-glucoproteína o puede estar mediada otros mecanismos, o puede suceder cuando se infecta a un mamífero con un microorganismo o una combinación de microorganismos resistentes a múltiples fármacos. La determinación de la resistencia a un tratamiento terapéutico es rutinaria en la técnica y se encuentra dentro de las capacidades de un profesional sanitario experto habitual, por ejemplo, puede medirse mediante ensayos de proliferación celular y ensayos de muerte celular, tal como se describe anteriormente en el presente documento para "sensibilización".

Tal como se usa en el presente documento, el término "quimioterapia" se refiere al uso de fármacos para destruir células cancerosas (incluyendo leucemias y linfomas). Hay más de 50 fármacos de quimioterapia diferentes y algunos se administran solos, pero a menudo pueden combinarse varios fármacos (conocido como quimioterapia combinada). La quimioterapia puede usarse sola para tratar algunos tipos de cáncer. En ocasiones puede usarse junto con otros tipos de tratamiento, tales como cirugía, radioterapia, inmunoterapia, o una combinación de los mismos.

Tal como se usa en el presente documento, "radioterapia", también llamada "terapia de radiación", se refiere al tratamiento del cáncer y de otras enfermedades con radiación ionizante. La radiación ionizante deposita energía de lesiona o destruye a las células en el área que se esté tratando (el "tejido diana") dañando a su material genético, haciendo imposible que estas células continúen creciendo. Aunque la radiación daña a células tanto cancerosas como normales, las últimas son capaces de repararse a sí mismas y funcionar de manera adecuada. La radioterapia puede usarse para tratar tumores sólidos localizados, tales como cánceres de piel, de lengua, de laringe, de cerebro, de mama, o de cuello de útero. También puede usarse para tratar leucemias y linfomas (cánceres de las células formadoras de la sangre y del sistema linfático, respectivamente).

Tal como se usa en el presente documento, la expresión "protocolo de tratamiento" se refiere al proceso de informar de la toma de decisiones para el tratamiento de una enfermedad. Tal como se usa en el presente documento, el protocolo de tratamiento se basa en los niveles comparativos de uno o más polipéptidos relacionados con el crecimiento celular en una muestra de tejido de un paciente en relación a los niveles del mismo polipéptido (o polipéptidos) en una diversidad de muestras de tejido normal y enfermo para el que hay disponible información del paciente, incluyendo las estrategias de tratamiento y los resultados.

Tal como se usa en el presente documento, la expresión "características biológicas" se refiere al fenotipo y/o genotipo de una o más células o tejidos, que pueden incluir el tipo celular, y/o el tipo tisular a partir del cual se obtuvo la célula, las características morfológicas de la célula (o células)/tejido (o tejidos), y la expresión de moléculas biológicas en la célula (o células)/tejido (o tejidos).

Tal como se usa en el presente documento, el término "muestra" se refiere a material procedente del cuerpo de un animal, incluyendo, pero sin limitación, sangre, suero, el plasma, orina, aspirado del pezón, fluido cefalorraquídeo, hígado, riñón, de mama, hueso, médula ósea, testículos u ovarios y cerebro, de colon, y pulmón. Una "muestra" de acuerdo con la presente invención, también puede ser células y tejidos cultivados.

Tal como se usa en el presente documento, una "muestra de cáncer" se refiere a una muestra que se origina a partir de un cáncer, es decir, a partir de un nuevo crecimiento de tejido diferente o anormal. Una "muestra de cáncer" puede ser una muestra celular o de tejido. Las células de cáncer pueden existir como parte del tejido de cáncer, o puede existir como células que flotan libremente desprendidas del tejido canceroso a partir del cual se originaron. Una muestra de cáncer, de acuerdo con la presente invención, puede usarse para pruebas de cánceres *in vitro* o *ex vivo*.

Tal como se usa en el presente documento, la expresión "muestra no de cáncer" se refiere a una muestra de células o tejido obtenida de un tejido normal. Un experto en la técnica puede juzgar que una muestra no sea una muestra tumoral basándose en la morfología y otras pruebas diagnósticas.

Tal como se usa en el presente documento, el término "mamífero" se refiere a un ser humano u otro animal, tales como animales de granja o animales de laboratorio (por ejemplo, cobayas o ratones).

Tal como se usa en el presente documento, "hibridación específica" o "hibridación selectiva" o "hibridación en condiciones rigurosas" se refiere a la hibridación que tiene lugar cuando dos secuencias polinucleotídicas son sustancialmente complementarias, es decir, que hay al menos aproximadamente un 60% y preferentemente, al menos aproximadamente un 80% o 90% de identidad entre las secuencias, en el que la región de identidad

comprende al menos 10 nucleótidos. En una realización, las secuencias hibridan en condiciones rigurosas después de la incubación de las secuencias durante toda una noche a 42°C, seguido de lavados rigurosos (SSC 0,2X a 65°C). Típicamente, las condiciones rigurosas serán aquellas en las que la concentración de sal sea al menos de 0,01 a 1,0 M de concentración de ión Na (u otras sales) a pH 7,0 a 8,3 y la temperatura sea de al menos

aproximadamente 30°C para sondas cortas (por ejemplo, de aproximadamente 6 a 50 nucleótidos). También pueden lograrse condiciones rigurosas mediante la adición de agentes desestabilizantes, tales como formamida. En general, las condiciones rigurosas se seleccionan para que sean aproximadamente 5°C menores que el punto térmico de fusión (T_m) para la secuencia específica a una fuerza iónica y pH definidos, según se calcula usando métodos rutinarios en la técnica.

Tal como se usa en el presente documento, el término "homología" se refiere al alineamiento óptimo de secuencias (ya sean de nucleótidos o aminoácidos), que puede llevarse a cabo mediante implementaciones de algoritmos computarizadas. La "homología", con respecto a polinucleótidos, por ejemplo, puede determinarse mediante análisis con BLASTN, versión 2.0, usando los parámetros por defecto. La "homología", con respecto a polipéptidos (es decir, aminoácidos), puede determinarse usando un programa, tal como BLASTP, versión 2.2.2 con los parámetros por defecto, que alinea a los polipéptidos o fragmentos que se estén comparando y determina el alcance de identidad o similitud de aminoácidos entre ellos. Se apreciará que la "homología" de aminoácidos incluye sustituciones conservativas, es decir, aquellas que sustituyen a un aminoácido dado en un polipéptido por otro aminoácido de características similares. Los siguientes reemplazos se ven típicamente como sustituciones conservativas: reemplazos de un aminoácido alifático, tal como Ala, Val, Leu e Ile por otro aminoácido alifático; reemplazo de una Ser con un Thr o *vice versa*; reemplazo de un resto ácido, tal como Asp o Glu por otro resto ácido; reemplazo de un resto que porta un grupo amida, tal como Asn o Gln, por otro resto que porta un grupo amida; intercambio de un resto básico, tal como Lys o Arg con otro resto básico; y reemplazo de un resto aromático, tal como Phe o Tyr por otro resto aromático. Una "homología del 60% o mayor" incluye una homología de, por ejemplo, 60%, 70%, 75%, 80%, 85%, 90%, 95%, y hasta el 100% (idéntica) entre dos o más secuencias de nucleótidos o de aminoácidos.

Tal como se usa en el presente documento, el término "polinucleótido" incluye ARN, ADNc, ADN genómico, formas sintéticas, y polímeros mixtos, hebras tanto con sentido como antisentido, y pueden modificarse química o bioquímicamente o pueden contener nucleobases no naturales o derivatizadas, tal como apreciarán fácilmente los expertos en la materia. Dichas modificaciones incluyen, por ejemplo, marcadores, metilación, sustitución de uno o más de los nucleótidos de origen natural con un análogo, modificaciones internucleótido, tales como enlaces cargados (por ejemplo, fosfonatos de metilo, fosfotriésteres, fosfoamidatos, carbamatos, etc.), enlaces cargados (por ejemplo, fosforotioatos, fosforoditioatos, etc.), restos colgantes (por ejemplo, polipéptidos), y enlaces modificados (por ejemplo, polinucleótidos alfa anómicos, etc.). También se incluyen moléculas sintéticas que imiten a los polinucleótidos en su capacidad para unirse a una secuencia designada mediante enlace de hidrógeno y otras interacciones químicas.

Tal como se usa en el presente documento, el término "mutación" se refiere a un cambio en una secuencia nucleotídica dentro de un gen, o fuera del gen en una secuencia reguladora, en comparación con el tipo silvestre. El cambio puede ser una eliminación, sustitución, mutación puntual, mutación de múltiples nucleótidos, transposición, inversión, desplazamiento de marco, mutación sin sentido u otras formas de aberración que diferencien al polinucleótido o secuencia de proteína de la de un gen expresado de manera normal en una célula funcional donde la expresión y la funcionalidad se encuentran dentro del intervalo de origen natural.

"Polipéptido" y "proteína" se usan de manera intercambiable en la presente memoria para referirse a un polímero de restos de aminoácidos. La expresión "proteína recombinante" se refiere a una proteína que se produce mediante la expresión de una molécula de ADN recombinante que codifica la secuencia de aminoácidos de la proteína. Los polinucleótidos y polipéptidos producidos de manera recombinante, y fragmentos o análogos de los mismos, pueden prepararse de acuerdo con métodos conocidos en la técnica, y descritos en Maniatis *et al.*, Molecular Cloning: A Laboratory Manual, 2ª Ed., (1989), Cold Spring Harbor, N.Y., y Berger y Kimmel, Methods in Enzymology, Volumen 152, Guide to Molecular Cloning Techniques (1987), Academic Press, Inc., San Diego, Calif..

Tal como se usa en el presente documento, el término "fragmento", cuando va en referencia a un polipéptido (como en "un fragmento de una proteína dada") se refiere a una parte más corta del polipéptido. El fragmento puede tener un tamaño que varía desde cuatro restos de aminoácidos hasta la secuencia de aminoácidos completa menos un aminoácido. En una realización, la presente invención contempla "fragmentos funcionales" de un polipéptido. Dichos fragmentos, tal como se enseña en el presente documento, son "funcionales" en tanto que retengan la capacidad para sensibilizar a una muestra de cáncer o a un mamífero con cáncer a un tratamiento terapéutico, aunque quizá con una actividad sensibilizante menor que la observada para el polipéptido de longitud completa. Dicho "fragmento" de un polipéptido es preferentemente mayor de 10 aminoácidos de longitud, y más preferentemente mayor de 50 aminoácidos de longitud, y aún más preferentemente mayor de 100 aminoácidos de longitud. Un "fragmento" de un polipéptido de la familia de SPARC, de acuerdo con la presente invención, puede contener uno o más de los tres dominios conservados de los miembros de la familia de SPARC, es decir, el dominio N-terminal ácido, el dominio similar a folistatina, y el dominio de unión a calcio extracelular EC.

Tal como se usa en el presente documento, una "variante" de un polipéptido específico se refiere a un polipéptido sustancialmente similar al péptido específico concreto o a un fragmento del mismo. Por "sustancialmente similar", se entiende que se hace que la variante llegue a una construcción final que posea la función deseada, es decir, sensibilizar a una muestra de cáncer o a un mamífero a un tratamiento terapéutico, aunque quizá con una menor actividad sensibilizante de la variante que la observada para el polipéptido de tipo silvestre. Dichas variantes incluyen, por ejemplo, eliminaciones, inserciones, o sustituciones de restos dentro de la secuencia de aminoácidos del polipéptido específico. Además, una "variante" también puede ser un polipéptido de fusión entre un polipéptido de la familia de SPARC y un segundo polipéptido. Puede efectuarse cualquier combinación de eliminación, inserción, sustitución y fusión.

Tal como se usa en el presente documento, "aislado" o "purificado", cuando se usan en referencia a un polinucleótido o polipéptido significan que se ha retirado una secuencia de nucleótidos o de aminoácidos de origen natural de su ambiente celular normal o que se sintetiza en un ambiente no natural (por ejemplo, se sintetiza artificialmente). Por lo tanto, una secuencia "aislada" o "purificada" puede estar en una solución sin células o ponerse en un ambiente celular diferente. El término "purificado" no implica que la secuencia de nucleótidos o aminoácidos sea el único polinucleótido o polipéptido presente, pero que esté esencialmente libre (aproximadamente el 90-95%, hasta el 99-100% de pureza) de material no polinucleotídico o polipeptídico naturalmente asociado con este.

Tal como se usa en el presente documento, el término "codifica" se refiere a la propiedad inherente de las secuencias de nucleótidos específicas en un polinucleótido, tal como un gen en un cromosoma o un ARNm, para servir como moldes para la síntesis de otros polímeros y macromoléculas en procesos biológicos que tienen una secuencia de nucleótidos definida (es decir, ARNr, ARNt, otras moléculas de ARN) o de aminoácidos y las propiedades biológicas que resulten de las mismas. Por lo tanto, un gen codifica a una proteína, si la transcripción y la traducción de ARNm producido por ese gen producen la proteína en una célula u otro sistema biológico. Tanto la hebra codificante, la secuencia nucleotídica que es idéntica a la secuencia de ARNm y se proporciona normalmente en los listados de secuencias, como la hebra no codificante, usada como molde para la transcripción, de un gen o ADNc puede citarse como codificante de la proteína u otro producto de ese gen o ADNc. Un polinucleótido que codifica una proteína incluye cualquier polinucleótido que tenga secuencias nucleotídicas diferentes pero que codifiquen la misma secuencia de aminoácidos de la proteína debido a la degeneración del código genético. Los polinucleótidos y secuencias nucleotídicas que codifican proteínas pueden incluir intrones y pueden ser ADN genómico.

Tal como se usa en el presente documento, el término "vector" se refiere a un compuesto polinucleotídico usado para introducir polinucleótidos exógenos o endógenos en células hospedadoras. Un vector comprende una secuencia nucleotídica que puede codificar a una o más moléculas de polipéptido. Los plásmidos, cósmidos, virus y bacteriófagos, en un estado natural o que hayan sido sometidos a ingeniería genética recombinante, son ejemplos no limitantes de vectores usados de manera común para proporcionar vectores recombinantes que comprenden al menos una molécula de polinucleótido aislada deseada.

Tal como se usa en el presente documento, el término "transformación" o el término "transfección" se refieren a una diversidad de técnicas reconocidas en la materia para introducir un polinucleótido exógeno (por ejemplo, ADN) en una célula. Una célula se "transforma" o "transfecta" cuando se ha introducido ADN exógeno en el interior de la membrana celular. Los términos "transformación" y "transfección" y los términos provenientes de cada uno se usan de manera indistinta.

Tal como se usa en el presente documento, un "vector de expresión" se refiere a un casete de expresión recombinante que tiene un polinucleótido que codifica a un polipéptido (es decir, una proteína) que puede transcribirse y traducirse por una célula. El vector de expresión puede ser un plásmido, virus, o fragmento polinucleotídico.

El término "expresión" se refiere a la producción de una proteína o secuencia nucleotídica en una célula o en un sistema sin células, e incluye la transcripción en un producto de ARN, la modificación postranscripcional y/o la traducción en un producto de proteína o polipéptido a partir del ADN que codifica ese producto, así como las posibles modificaciones postraduccionales.

Tal como se usa en el presente documento, la expresión "comparar el nivel de expresión" se refiere a comparar la expresión diferencial de un polinucleótido o un polipéptido en dos o más muestras.

Tal como se usa en el presente documento, la expresión "expresión diferencial" se refiere a diferencias tanto cuantitativas como cualitativas en los patrones de expresión en un polinucleótido o un polipéptido entre dos o más muestras. Se dice que un polinucleótido o un polipéptido se "expresa diferencialmente" si su expresión es detectable en una muestra, pero no en otra muestra, mediante métodos conocidos para la detección de polinucleótidos o polipéptidos (por ejemplo, electroforesis). También se dice que un polinucleótido o un polipéptido se "expresa diferencialmente" si la diferencia cuantitativa de su expresión (es decir, aumento o disminución, medido en µg, µmol o número de copias) entre dos muestras es de aproximadamente el 20%, aproximadamente el 30%,

aproximadamente el 50%, aproximadamente el 70%, aproximadamente el 90% a aproximadamente el 100% (aproximadamente el doble) o más, hasta incluyendo aproximadamente hasta 1,2 veces, 2,5 veces, 5 veces, 10 veces, 20 veces, 50 veces o más. Un transcrito génico "expresado diferencialmente" significa un transcrito de ARNm que se encuentra en diferentes números de copias entre dos o más muestras.

Tal como se usa en el presente documento, una "cantidad aumentada de polipéptido o polinucleótido de la familia de SPARC" se refiere a un nivel de expresión mayor de al menos el 20%, por ejemplo, 30%, 40%, 50%, 60%, 70%, 80%, o 90%, o 2 veces, 3 veces, 4 veces, 5 veces o más en una célula, en comparación con una célula de control. Se dice que una célula que expresa un polipéptido o polinucleótido de la familia de SPARC tiene una cantidad aumentada de dicho polipéptido o polinucleótido si la expresión es tal como se define anteriormente en el presente documento en comparación con una célula resistente a la quimioterapia.

La expresión "proteína secretada" se refiere a una proteína que tiene al menos una parte que es extracelular e incluye proteínas que son completamente extracelulares (es decir, no unidas a una célula). El "nivel de secreción" se refiere al nivel (es decir, cantidad) de una proteína secretada en el compartimento extracelular.

Tal como se usa en el presente documento, el término "proliferación" se refiere a la velocidad de la división celular y a la capacidad de una célula para continuar dividiéndose. Un proceso de división celular completo se cita como un "ciclo". Por un "aumento en la proliferación celular" se entiende un aumento en la velocidad de división celular de tal forma que la célula tiene una mayor velocidad de división celular en comparación con células normales de ese tipo celular, o permitir que la división celular continúe durante más ciclos sin cambiar la velocidad de cada división celular, por ejemplo, un aumento del 10% o mayor (por ejemplo, 20%, 30%, 40%, 50%, hasta 2 veces, 5 veces, 10 veces o mayor). Por una "disminución en la proliferación celular" se entiende una disminución en la velocidad de división celular de tal forma que la célula tiene una menor velocidad de división celular en comparación con células normales de ese tipo celular, o reducir el número de ciclos de la división celular sin cambiar la velocidad de cada división celular, por ejemplo, disminución de un 10% o mayor (por ejemplo, 20%, 30%, 40%, 50%, hasta 2 veces, 5 veces, 10 veces o mayor).

La presente invención proporciona un polinucleótido de la familia de SPARC que codifica un polipéptido de la familia de SPARC para sensibilizar para los tratamientos terapéuticos para el cáncer. Dichas composiciones y métodos sensibilizantes son particularmente útiles para potenciar la respuesta de los pacientes que son resistentes a un tratamiento. También son útiles para reducir los efectos secundarios de la terapia para el cáncer, por ejemplo, potenciando la respuesta de un paciente a una fuerza menor (es decir, dosificación) del tratamiento. La composición de la presente invención puede reducir la dosificación de un agente de tratamiento terapéutico en al menos un 20%, por ejemplo, 30%, 40%, 50%, y hasta un 60%.

Los polipéptidos y polinucleótidos de la familia de SPARC

La presente divulgación enseña (a) un polipéptido de la familia de SPARC seleccionado entre el grupo que consiste en las SEC ID N°: 1-17; (b) un polipéptido que tiene una secuencia de aminoácidos con al menos un 60% de homología con dicho polipéptido de la familia de SPARC de (a); (c) un fragmento polipeptídico de (a)-(b) en el que dicho fragmento tiene al menos 50 aminoácidos de longitud, y (d) un polipéptido de fusión que comprende el polipéptido de (a)-(c).

Un polipéptido de la familia de SPARC puede ser un polipéptido de tipo silvestre o una variante del mismo, puede ser el polipéptido de longitud completa o un fragmento del mismo.

Basándose en la homología de secuencia, se han identificado varios miembros de la familia de SPARC, tales como SMOG-1 (Vannahme *et al.*, 2002, J. Biol. Chem., 277(41):37977-37986), hevina (Bendik I, Schraml P, Ludwig CU, Cancer Res. 1998; 58(4):626-9), SC1 (Johnston IG, Paladino T, Gurd JW, Brown IR, Neuron. 1990 4(1):165-76), QR-1, proteínas similares a folistatina (TSC-36) (Shibanuma, M., Mashimo, J., Mita, A., Kuroki, T. y Nose, K, 1993, Eur. J. Biochem. 217 (1) 13-19) y testicano (Alliel, P.M., Perin, J.P., Jolles, P. y Bonnet, F.J, 1993, Eur. J. Biochem. 214 (1), 347-350). Un polipéptido de la familia de SPARC incluye, pero sin limitación, SPARC y cualquiera de sus miembros de la familia identificados conocidos en la técnica o tal como se describen en el presente documento.

El alineamiento de las diversas secuencias de polinucleótido permiten a un experto en la materia seleccionar partes conservadas de las proteínas (es decir, aquellas partes en común entre dos o más secuencias) así como partes no conservadas (es decir, aquellas partes únicas para dos o más secuencias).

Preferentemente, el polipéptido de la familia de SPARC de la presente invención contiene uno o dos o tres dominios (es decir, el dominio ácido N-terminal, el dominio similar a folistatina, y/o el dominio EC), conservados entre los miembros de la familia de SPARC.

El fragmento sensibilizante a la terapia o una variante de una proteína de la familia de SPARC de tipo silvestre puede delinearse mediante manipulaciones de secuencia rutinarias conocidas para los expertos en la materia, incluyendo, pero sin limitación, mutaciones de eliminación, mutaciones puntuales, fusiones, tal como se describe

más adelante en el presente documento y tal como se describe en J. Sambrook, E. F. Fritsch, y T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2ª Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York, 1989, y Ausubel *et al.*, *Current Protocols in Molecular Biology*, 1994.

- 5 Una mutación en el ADN que codifica al polipéptido variante no debe alterar la fase de lectura y preferentemente no creará regiones complementarias que puedan producir estructuras secundarias en el ARNm. A nivel genético, estas variantes se preparan mediante mutagénesis de sitio dirigido de nucleótidos en el ADN que codifica la molécula peptídica, produciendo de este modo ADN que codifica la variante, y posteriormente expresar el ADN en cultivo celular recombinante.

10

Producción y purificación in vitro de un polipéptido de la familia de SPARC

- Puede producirse un polipéptido de la familia de SPAR en una célula procariota o eucariota usando cualquier método conocido, por ejemplo, técnicas de ADN recombinante. Las técnicas de recombinación pueden llevarse a cabo tal como se describe a continuación en el presente documento, o por ejemplo, mediante los métodos divulgados en T. Maniatis *et al.*, *"Molecular Cloning"*, 2ª Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N. T. (1989); Nippon Seikagaku Kai (Biochemical Society of Japan) ed., *"Zoku-Seikagaku Jikken Kouza 1, Idenshi Kenkyuho II (Lectures on Biochemical Experiments (Segunda serie; 1), Methods for Gene Study II)"*, Tokyo Kagaku Dojin, Japón (1986); Nippon Seikagaku Kai (Biochemical Society of Japan) ed., *"Shin-Seikagaku Jikken Kouza 2, Kakusan III (Kumikae DNA Gijutsu) (New Lectures on Biochemical Experiments 2, Nucleic Acids III (Recombinant DNA Technique))"*, Tokyo Kagaku Dojin, Japón (1992); R. Wu (ed.), *"Methods in Enzymology"*, Vol. 68, Academic Press, Nueva York (1980); R. Wu *et al.* (ed.), *"Methods in Enzymology"*, volúmenes 100 y 101, Academic Press, Nueva York (1983); R. Wu *et al.* (ed.), *"Methods in Enzymology"*, Volúmenes 153, 154 y 155, Academic Press, Nueva York (1987), etc., así como mediante técnicas divulgadas en las referencias citadas en los mismos. Dichas técnicas y medios también pueden ser aquellas que se modifican/mejoran individualmente a partir de técnicas convencionales, dependiendo del objeto de la presente invención.

- Un polipéptido de la familia de SPARC puede expresarse y purificarse a partir de una célula hospedadora recombinante. Las células hospedadoras recombinantes pueden ser procariotas o eucariotas, incluyendo, sin limitación, bacterias, tales como *E. coli*, células fúngicas, tales como levaduras, células de insecto, incluyendo, pero sin limitación líneas celulares procedentes de *Drosophila* y de gusano de seda, y células y líneas celulares de mamífero.

- En determinadas realizaciones, cuando se expresa y purifica un polipéptido de la familia de SPARC de la presente invención, se emplean técnicas para mejorar la solubilidad de proteínas para prevenir la formación de cuerpos de inclusión (que son fracciones insolubles), y por lo tanto obtener grandes cantidades del polipéptido. SPARC acumulado en cuerpos de inclusión es un SPARC de tipo inactivo que no mantiene sus propiedades fisiológicas.

- La solubilidad de un polipéptido de la familia de SPARC purificado puede mejorarse mediante métodos conocidos en la técnica, y tal como se describen a continuación en el presente documento.

- Por ejemplo, también puede mejorarse la solubilidad expresando un fragmento funcional, pero no el polipéptido de la familia de SPARC de longitud completa. El fragmento expresado debe mantener la actividad sensibilizante, tal como se ha descrito en el presente documento, aunque la actividad puede ser menor que la de un polipéptido de longitud completa.

- Puede expresarse un fragmento que contiene uno, dos o tres dominios de los tres dominios conservados de los miembros de la familia de SPARC.

- Además, para aumentar la solubilidad de una proteína expresada (por ejemplo, en *E. coli*), se puede reducir la velocidad de síntesis de la proteína reduciendo la temperatura de cultivo, usando un promotor más débil, usando un plásmido con menor número de copias, disminuyendo la concentración de inductor, cambiando el medio de cultivo, tal como se describe en Georgiou, G. y Valax, P. (1996, *Current Opinion Biotechnol.* 7, 190-197). Esto disminuye la velocidad de síntesis de proteínas y normalmente se obtiene más proteína soluble. También pueden usarse grupos prostéticos o cofactores que sean esenciales para el plegamiento adecuado o para la estabilidad de la proteína, o añadirse un tampón para controlar la fluctuación del pH en el medio durante el cultivo, o añadir glucosa al 1% para suprimir la inducción del promotor *lac* por lactosa, que está presente en la mayoría de medios ricos (tales como LB, 2xYT). También pueden añadirse polioles (por ejemplo, sorbitol) al medio porque el aumento en la presión osmótica causada por estas adiciones da lugar a la acumulación de osmoprotectores en la célula, que estabilizan la estructura de la proteína nativa. Pueden añadirse etanol, tioles de bajo peso molecular y disulfuros y NaCl. Además, pueden expresarse conjuntamente con el polipéptido deseado chaperonas y/o foldasas. Las chaperonas moleculares promueven la isomerización adecuada y el direccionamiento celular interactuando de manera transitoria con intermedios del plegamiento. Los sistemas de chaperona de *E. coli* mejor caracterizados son: GroES-GroEL, DnaK-DnaJ-GrpE, C1pB. Las foldasas aceleran las etapas limitantes de la velocidad durante la ruta del plegamiento. Tres tipos de foldasas desempeñan un papel importante: peptidil prolil cis/trans isomerasas (PPI), disulfuro oxidoreductasa (DsbA) y disulfuro isomerasa (DsbC), proteína disulfuro isomerasa (PDI) que es una proteína de

eucariota que cataliza tanto la oxidación de cisteínas de la proteína como la isomerización del enlace disulfuro. La coexpresión de una o más de estas proteínas con la proteína diana podría dar lugar a mayores niveles de proteína soluble.

Puede producirse un polipéptido de la familia de SPARC como una proteína de fusión para mejorar su solubilidad y producción. La proteína de fusión comprende un polipéptido de la familia de SPARC y un segundo polipéptido fusionados juntos en fase. El segundo polipéptido puede ser un compañero de fusión conocido en la técnica por mejorar la solubilidad del polipéptido al que se fusiona, por ejemplo, NusA, bacterioferritina (BFR), GrpE, tioredoxina (TRX) y glutatión-S-transferasa (GST). Novagen Inc., radicada en Madison, Wis., proporciona la serie de vectores pET43.1 que permite la formación de una fusión de NusA-diana. DsbA y DsbC también han mostrado efectos positivos en los niveles de expresión cuando se usan como compañeros de fusión, y por lo tanto pueden usarse para fusionarse con un polipéptido de SPARC para lograr una mayor solubilidad.

Tal como se divulga en el presente documento, se produce un polipéptido de la familia de SPARC como polipéptido de fusión que comprende el polipéptido de la familia de SPARC y una tioredoxina compañera de fusión, tal como se describe en la patente de Estados Unidos n.º 6.387.664. La fusión de tioredoxina-SPARC puede producirse en *E. coli* como una proteína soluble fácil de formular en grandes cantidades sin perder las actividades fisiológicas de SPARC. Aunque la Patente de Estados Unidos n.º 6.387.664 proporciona una proteína de fusión de SPARC con SPARC fusionado al extremo C-terminal de la tioredoxina, se entiende, para el fin de la presente invención, que puede fusionarse un polipéptido de la familia de SPARC al extremo N-terminal o al extremo C-terminal de un segundo polipéptido, siempre que se mantenga su función sensibilizante.

Además de aumentar la solubilidad, puede construirse una proteína de fusión que comprenda un polipéptido de la familia de SPARC para la fácil detección de la expresión del polipéptido de la familia de SPARC en una célula. En una realización, el segundo polipéptido que está fusionado al polipéptido de la familia de SPARC es un polipéptido indicador. El polipéptido indicador, cuando sirve para dicho fin de detección, no tiene que fusionarse con el polipéptido de la familia de SPARC. Puede estar codificado por el mismo polinucleótido (por ejemplo, un vector) que también codifica al polipéptido de la familia de SPARC e introducirse y expresarse conjuntamente en una célula diana.

El polipéptido indicador, tal como se enseña en el presente documento puede ser una proteína autofluorescente (por ejemplo, GFP, EGFP). Las proteínas autofluorescentes proporcionan un ensayo listo para la identificación de la expresión de un polinucleótido (y del producto polipeptídico) de interés. Debido a que la actividad del polipéptido indicador (y por inferencia su nivel de expresión) puede controlarse de manera cuantitativa usando un clasificador celular, es sencillo ensayar muchos transfectantes independientes ya sea de manera secuencial o en la población en bruto. Pueden seleccionarse de entre la población las células con la mejor expresión. Esto es útil cuando se selecciona una célula recombinante que comprende un polipéptido o un polinucleótido de la familia de SPARC para sensibilizar para el tratamiento de acuerdo con la presente invención. Los parámetros cuantitativos, tales como la intensidad media de fluorescencia y la varianza pueden determinarse a partir del perfil de intensidad de fluorescencia de la población celular (Shapiro, H., 1995, Practical Flow Cytometry, 217-228). Los ejemplos no limitantes de moléculas indicadoras útiles en la invención incluyen luciferasa (de luciérnaga u otras especies), cloranfenicol acetiltransferasa, β - galactosidasa, proteína fluorescente verde (GFP), proteína fluorescente verde potenciada (EGFP), y dsRed.

La expresión del polipéptido de SPARC (ya sea solo, o como una proteína de fusión) también puede determinarse directamente mediante un inmunoensayo, tal como un ELISA (ensayo inmunoabsorbente ligado a enzimas) (véanse, por ejemplo, la Patente de Estados Unidos n.º 5.962.320; la Patente de Estados Unidos n.º 6.187.307; la Patente de Estados Unidos n.º 6.194.205), transferencia de Western, o mediante otros métodos rutinarios en la técnica. La expresión de un polipéptido de la familia de SPARC puede detectarse indirectamente detectando el transcrito de la proteína (por ejemplo, mediante análisis de hibridación, tal como transferencia de Northern o micromatriz de ADN, o mediante PCR).

En una realización, se fusiona un polinucleótido que codifica un segundo polipéptido a un polinucleótido que codifica un polipéptido de la familia de SPARC mediante una secuencia enlazadora intermedia que codifica un enlazador polipeptídico.

En otra realización, el polipéptido enlazador comprende un sitio de escisión de proteasa que comprende un enlace peptídico que puede hidrolizarse con una proteasa. Como resultado, el polipéptido de la familia de SPARC puede separarse del segundo polipéptido tras su expresión mediante proteólisis. El enlazador puede comprender uno o más aminoácidos adicionales a ambos lados del enlace al que también se une el sitio catalítico de la proteasa (véase, por ejemplo, Schechter y Berger, 1967, Biochem. Biophys. Res. Commun. 27, 157-62). Como alternativa, el sitio de escisión del enlazador puede separarse del sitio de reconocimiento de la proteasa y pueden separarse el sitio de escisión y el sitio de reconocimiento por uno o más (por ejemplo, de dos a cuatro) aminoácidos. En un aspecto, el enlazador comprende al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50 o más aminoácidos. Más preferentemente, el enlazador tiene entre 5 y 25 aminoácidos de longitud, y lo más preferentemente, el enlazador tiene entre 8 y 15 aminoácidos de longitud.

5 Algunas proteasas útiles de acuerdo con la invención se discuten en las siguientes referencias: V.Y.H. Hook, Proteolytic and cellular mechanisms in prohormone and proprotein processing, RG Landes Company, Austin, Texas, EE.UU. (1998); N.M. Hooper *et al.*, 1997, Biochem. J. 321:265-279; Werb, 1997, Cell 91: 439-442; Wolfsberg *et al.*, 1995, J. Cell Biol. 131: 275-278; Murakami y Etlinger, 1987, Biochem. Biophys. Res. Comm. 146: 1249-1259; Berg *et al.*, 1995, Biochem. J. 307: 313-326; Smyth y Trapani, 1995, Immunology Today 16: 202-206; Talanian *et al.*, 1997, J. Biol. Chem. 272: 9677-9682; y Thornberry *et al.*, 1997, J. Biol. Chem. 272: 17907-17911. Las proteasas adecuadas incluyen, pero sin limitación, aquellas listadas en la tabla 1 a continuación.

TABLA 1. Proteasas y sus señales de escisión	
Proteasa	Señal de escisión (Secuencia polinucleotídica enlazadora ejemplar)
familia de subtilisina/kexina (furina, PC1, PC2, PC4, PACE4, PC5, PC)	RXKR (SEC ID N°: 49). (CGC NNN AAG CGC) (SEC ID N°: 50).
MMP-2	PLGLWA (SEC ID N°: 51). (CCC CTG GGC CTG TGG GCC) (SEC ID N°: 52).
MT1-MMP	PLGLWA (SEC ID N°: 51). (CCC CTG GGC CTG TGG GCC) (SEC ID N°: 52).
Proteasa	Señal de escisión - Secuencia de aminoácidos (Secuencia polinucleotídica enlazadora ejemplar)
caspasa-1	YEYD GW (SEC ID N°: 53). (TAC GAG GTG GAC GGC TGG) (SEC ID N°: 54).
caspasa-2	VDVAD GW (SEC ID N°: 55). (GTG GAC GTG GCC GAC GGC TGG) (SEC ID N°: 56)
caspasa-3	VDQMD GW (SEC ID N°: 57). (GTG GAC CAG ATG GAC GGC TGG) (SEC ID N°: 58).
caspasa-4	LEV DGW (SEC ID N°: 59). (CTG GAG GTG GAC GGC TGG) (SEC ID N°: 60).
caspasa-6	VQV DGW (SEC ID N°: 61). (GTG CAG GTG GAC GGC TGG) (SEC ID N°: 62).
caspasa-7	VDQV DGW (SEC ID N°: 63). (GTG GAC CAG GTG GAC GGC TGG) (SEC ID N°: 64).
caspasa-8	DXXD (SEC ID N°: 65). (GAC NNN NNN GAC) (SEC ID N°: 66).

TABLA 1. Proteasas y sus señales de escisión	
Proteasa	Señal de escisión (Secuencia polinucleotídica enlazadora ejemplar)
caspasa-9	DXXD (SEC ID N°: 65). (GAC NNN NNN GAC)
alfa-secretasa	proteína precursora de amiloide (APP)
convertasa de proproteína (subtilisina/kexina isozima SKI-1)	RGLT (SEC ID N°: 67). (CGC GGC CTG ACC) (SEC ID N°: 68).
convertasas de proproteína	escisión en restos hidrófobos (por ejemplo, Leu, Phe, Val, o Met) o en restos de aminoácidos pequeños, tales como Ala o Thr
virus de la fiebre aftosa, proteasa 2A	NFDLLKLAGDVESNPGP (SEC ID N°: 69). (AAC TTC GAC CTG CTG AAG CTG GCC GGC GAC GTG GAG AGC AAC CCC GGC CCC) (SEC ID N°: 70)
peptidasa de señal	A-X-A-X (SEC ID N°: 71). (GCC NNN GCC NNN) (SEC ID N°: 72).
aminopeptidasas (por ejemplo, arginina aminopeptidasa, lisina aminopeptidasa, aminopeptidasa B, tripsina)	LTK (SEC ID N°: 73). (CTG ACC AAG) (SEC ID N°: 74).
enzima degradante de insulina	GGFLRKVGQ (SEC ID N°: 75). (GGC GGC TTC CTG CGC AAG GTG GGC CAG) SEC ID N°: 76)

Pueden obtenerse péptidos enlazadores adicionales a partir de los sustratos para la enzima convertidora de la proopiomelanocortina (PCE); proteasa aspártica de gránulos cromafines (CGAP); tiol proteasa de prohormonas; carboxipeptidasas (por ejemplo, carboxipeptidasa E/H, carboxipeptidasa D y carboxipeptidasa Z); pirrolilendopeptidasas; y proteasa de alto peso molecular.

Las proteasas de la superficie celular también pueden usarse con enlazadores escindibles de acuerdo con la invención e incluyen, pero sin limitación: aminopeptidasa N; aminopeptidasa sensible a puomicina; enzima convertidora de angiotensina; proglutamil peptidasa II; dipeptidil peptidasa IV; convertasa dibásica de N-arginina; endopeptidasa 24.15; endopeptidasa 24.16; alfa, beta y gamma secretasas de proteína precursora de amiloide; secretasa de enzima convertidora de angiotensina; secretasa de TGF alfa; secretasa de TNF alfa; secretasa de ligando FAS; secretasas del receptor I y II de TNF; secretasa de CD30; secretasas de KL1 y KL2; secretasa del receptor de IL-6; secretasa de CD43, CD44; secretasas de CD16-I y CD16-II; secretasa de L-selectina; secretasa de receptor de folato; MMP 1, 2, 3, 7, 8, 9, 10, 11, 12, 13, 14, y 15; activador de plasminógeno de urocinasa; activador del plasminógeno tisular; plasmina; trombina; BMP-1 (procolágeno C-peptidasa); ADAM 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, y 11; y, granzimas A, B, C, D, E, F, G, y H.

Una alternativa a basarse en proteasas asociadas a células es el uso de un enlazador autoescindible. Por ejemplo, puede usarse como enlazador la proteasa 2A del virus de la fiebre aftosa (FMDV). Este es un polipéptido corto de 17 aminoácidos que escinde a la poliproteína de FMDV en la unión 2A/2B. La secuencia del propéptido de FMDV 2A es NFDLLKLAGDVESNPGP (SEC ID N°: 77). La escisión sucede en el extremo C-terminal del péptido en la pareja de aminoácidos glicina-prolina final y es independiente de la presencia de otras secuencias de FMDV y se escinde incluso en la presencia de secuencias heterólogas.

La inserción de esta secuencia entre dos regiones codificantes de proteínas (es decir, entre el polipéptido de la familia de SPARC y el segundo polipéptido de una proteína de fusión de acuerdo con la invención) da como

resultado la formación de una quimera autoescindible que se escinde en un fragmento C-terminal que porta la prolina C-terminal de la proteasa 2A en su extremo N-terminal, y un fragmento N-terminal que porta el resto del péptido de la proteasa 2A en su extremo C-terminal (véase, por ejemplo, P. de Felipe *et al.*, Gene Therapy 6: 198-208 (1999)). Los enlazadores autoescindibles y las combinaciones adicionales de proteasa-enlazador se describen adicionalmente en el documento WO 0120989.

Los polinucleótidos que codifican secuencias enlazadoras descritas anteriormente pueden clonarse a partir de secuencias que codifican los sustratos naturales de una proteasa adecuada o pueden sintetizarse químicamente usando métodos rutinarios en la técnica.

Cuando se expresa un polipéptido de la familia de SPARC en una célula humana, por ejemplo, *in vitro* o *in vivo*, los codones seleccionados para dicho polinucleótido que codifica al polipéptido de la familia de SPARC son preferentemente aquellos que se usen más frecuentemente en seres humanos, tales como aquellos listados en la tabla 2 a continuación. Las secuencias polinucleótídicas ejemplares mostradas en la tabla 4 se basan en codones que se usan con más frecuencia en seres humanos.

TABLA 2. Codones de ADN preferidos para uso en seres humanos			
Aminoácidos	Código de 3 letras	Código de 1 letra	Codones preferidos en genes humanos
Alanina	Ala	A	GCC GCT GCA GCG
Cisteína	Cys	C	TGT TGT
Ácido aspártico	Asp	D	GAC GAT
Ácido glutámico	Glu	E	GAG GAA
Fenilalanina	Phe	F	TTC TTT
Glicina	Gly	G	GGC GGG GGA GGT

TABLA 2. Codones de ADN preferidos para uso en seres humanos			
Aminoácidos	Código de 3 letras	Código de 1 letra	Codones preferidos en genes humanos
Histidina	His	H	CAC CAT
Isoleucina	Ile	I	ATC ATT ATA
Lisina	Lys	K	AAG AAA
Leucina	Leu	L	CTG TTG CTT <u>CTA</u> <u>TTA</u>
Metionina	Met	M	ATG
Asparagina	Asn	N	AAC AAT
Prolina	Pro	P	CCC CCT CCA CCG
Glutamina	Gln	Q	CAG CAA

TABLA 2. Codones de ADN preferidos para uso en seres humanos			
Aminoácidos	Código de 3 letras	Código de 1 letra	Codones preferidos en genes humanos
Arginina	Arg	R	CGC AGG CGG AGA CGA CGT
Serina	Ser	S	AGC TCC TCT AGT TCA TCG
Treonina	Thr	T	ACC ACA ACT ACG
Valina	Val	V	GTG GTC GTT GTA
Triptófano	Trp	W	TGG

TABLA 2. Codones de ADN preferidos para uso en seres humanos			
Aminoácidos	Código de 3 letras	Código de 1 letra	Codones preferidos en genes humanos
Tirosina	Tyr	Y	TAC TAT

Los codones situados en cabeza representan aquellos más preferidos para su uso en genes humanos. Los codones subrayados no se usan prácticamente nunca en genes humanos y por lo tanto no se prefieren.

- 5 También se divulga en el presente documento el uso de (a) un polinucleótido que codifica el polipéptido de un polipéptido de la familia de SPARC o un polipéptido de fusión que comprende un polipéptido de la familia de SPARC; y (b) un polinucleótido que hibrida con el polinucleótido de (a) en condiciones de hibridación rigurosas.

- 10 Las técnicas de manipulación de polinucleótidos para llevar a cabo las realizaciones anteriores de la invención se conocen bien en la técnica. (Véase, por ejemplo, Sambrook *et al.*, 1989; Ausubel *et al.* 1987 y en Annual Reviews of Biochemistry, 1992, 61:131-156). Los reactivos útiles para aplicar dichas técnicas, tales como enzimas de restricción y similares, se conocen ampliamente en la técnica y están disponibles comercialmente a través de una serie de comercializadores.

- 15 Las secuencias polinucleotídicas para su uso en la presente invención también pueden producirse totalmente o parcialmente mediante síntesis química, por ejemplo, mediante el método de fosoramidita descrito por Beaucage, *et al.*, 1981, Tetra. Letts., 22:1859-1862, o el método de triéster (Matteucci *et al.*, 1981, J. Am. Chem. Soc., 103:3185), que pueden efectuarse en sintetizadores de oligonucleótidos automatizados. Puede obtenerse un fragmento bicatenario a partir del producto monocatenario de síntesis química bien sintetizando la hebra complementaria e
20 hibridando las hebras juntas en condiciones adecuadas, o sintetizando la hebra complementaria usando ADN polimerasa con una secuencia cebadora adecuada.

En una realización, el polinucleótido de la familia de SPARC proporcionado por la presente invención existe en forma de vector, preferentemente, como vector de expresión.

25 Vectores de expresión

- Puede incorporarse un polinucleótido que codifica una secuencia polipeptídica de la familia de SPARC en vectores capaces de introducirse y replicarse en una célula procariota o eucariota para producir un polipéptido de la familia de
30 SPARC de la presente invención para un tratamiento de sensibilización, o puede usarse para transfectar o infectar a una célula o tejido y lograr la función de sensibilización directamente expresando el polipéptido de la familia de SPARC. Los vectores pueden integrarse o no en el genoma de la célula transfectada o infectada.

- 35 Las moléculas de polinucleótido útiles que codifican un polipéptido de la familia de SPARC de la presente invención pueden clonarse en un vector de expresión antes de que se introduzcan en una célula adecuada y pueden pasarse en células para generar cantidades usables de estos polinucleótidos.

- 40 Los vectores adecuados para la invención pueden ser vectores plasmídicos o virales. Los vectores de expresión plasmídicos incluyen, pero sin limitación, plásmidos, incluyendo pBR322, pUC, pcDNA3.1 o Bluescript™. Los vectores virales incluyen, pero sin limitación, baculovirus, adenovirus, poxvirus, virus adenoasociados (AAV), y vectores de retrovirus (Price *et al.*, 1987, Proc. Natl. Acad. Sci. USA, 84:156-160) tal como el vector pMV-7 incompetente para replicación basado en MMLV (Kirschmeier *et al.*, 1988, DNA, 7:219-225), así como cromosomas modificados humanos y de levadura (HAC y YAC).

- 45 Los vectores de expresión pueden comprender uno o más elementos reguladores para dirigir y/o potenciar la expresión de polinucleótidos cadena arriba o cadena abajo. Estas secuencias reguladoras se seleccionan basándose en las células que van a usarse para la introducción y/o expresión, y se unen operativamente a una secuencia polinucleotídica que va a expresarse. La expresión "elementos reguladores" pretende incluir promotores, potenciadores y otros elementos de control de la expresión (por ejemplo, señal de poliadenilación). Dichos
50 elementos reguladores se describen, por ejemplo, en Goeddel; 1990, Gene Expression Technology: Methods in Enzynology 185, Academic Press, San Diego, CA.

- Los elementos reguladores incluyen aquellos que dirigen la expresión de una secuencia nucleotídica en muchos tipos de células sujeto así como aquellos que dirigen la expresión de la secuencia nucleotídica únicamente en
55 determinadas células sujeto (por ejemplo, secuencias reguladoras específicas de tejido).

Los elementos reguladores incluyen también aquellos que dirigen la expresión constitutiva de una secuencia polinucleotídica unida operativamente y aquellos que dirigen la expresión inducible de la secuencia polinucleotídica.

Preferentemente, pueden usarse promotores adecuados. Por ejemplo, dichos promotores pueden incluir promotor de triptófano (trp), promotor de lactosa (lac), promotor de triptófano-lactosa (tac), promotor de lipoproteína (lpp), promotor de fago λ (P_L), etc. en los casos donde se usa *Escherichia coli* como hospedador; y los promotores GAL1 y GAL10, etc. en los casos donde se usa levadura como hospedador.

Algunos promotores y potenciadores de eucariotas tienen un intervalo amplio de células en las que pueden activar y/o modular la transcripción mientras que otros son solo funcionales en un subconjunto limitado de tipos celulares (véase, por ejemplo, Voss *et al.*, 1986, Trends Biochem. Sci., 11:287; y Maniatis *et al.*, anteriormente citado, para revisiones). Por ejemplo, el potenciador génico temprano de SV40 es muy activo en una gran variedad de tipos celulares de muchas especies de mamíferos y se ha usado ampliamente para la expresión de proteínas en células de mamífero (Dijkema *et al.*, 1985, EMBO J. 4:761). Otros dos ejemplos de elementos promotores/potenciadores activos en una gran variedad de tipos celulares de mamífero son aquellos del gen de factor de elongación (Uetsuki *et al.*, 1989, J. Biol. Chem., 264:5791; Kim *et al.*, 1990, Gene, 91:217; y Mizushima, *et al.*, 1990, Nagata, Nuc. Acids. Res., 18:5322) y las repeticiones terminales largas del virus del sarcoma de Rous (Gorman *et al.*, 1982, Proc. Natl. Acad. Sci. USA; 79:6777) y el citomegalovirus humano (Boshart *et al.*, 1985, Cell, 41:521).

Los promotores adecuados para expresión en células eucariotas incluyen, pero sin limitación, promotores TRAP, promotores adenovirales, tales como el promotor tardío principal adenoviral, el promotor de citomegalovirus (CMV); el promotor de virus respiratorio sincitial (RSV); promotores inducibles, tales como el promotor MMT, el promotor de metalotioneína, promotores de choque térmico; el promotor de albúmina; el promotor de ApoA1; promotores de globina humanos; promotores de timidina cinasa viral, tales como el promotor de timidina cinasa de *Herpes simplex*; LTR retrovirales; ITR; el promotor de β -actina; y promotores de hormona de crecimiento humana. El promotor también puede ser un promotor nativo que controle al polinucleótido que codifica al polipéptido y pueden encontrarse en la técnica las secuencias de promotores nativos (véase Agrawal *et al.*, 2000, J. Hematother. Stem Cell Res., 795-812; Courmoyer *et al.*, 1993, Annu. Rev. Immunol., 11:297-329; van de Stolpe *et al.*, 1996, J. Mol. Med., 74:13-33; Hermann, 1995, J. Mol. Med., 73:157-63)

Puede usarse una diversidad de secuencias potenciadoras en la presente invención, incluyendo, sin limitación: el potenciador de la cadena pesada de inmunoglobulina; el potenciador de la cadena ligera de inmunoglobulina; el potenciador de receptor de células T; los potenciadores DQ α y DQ β de HLA; el potenciador de β -interferón; el potenciador de interleucina 2; el potenciador del receptor de interleucina 2; el potenciador 5a^k de CMH de clase II; el potenciador de HLA-DR α de CMH de clase II; el potenciador de β -actina; el potenciador de creatina cinasa muscular; el potenciador de prealbúmina (transferrina); el potenciador de elastasa I; el potenciador de metalotioneína; el potenciador de colagenasa; el potenciador del gen de albúmina; el potenciador de α fetoproteína; el potenciador de β -globina; el potenciador de c-fos, el potenciador de c-HA-ras; el potenciador de insulina; el potenciador de la molécula de adhesión celular neural (NCAM); el potenciador de α -antitripsina; el potenciador de histonas H2B (TH2B); el potenciador de colágeno de ratón o de tipo I; el potenciador de proteínas reguladas por glucosa (GRP94 y GRP78); el potenciador de la hormona del crecimiento de ratas; el potenciador de amiloide A sérico (SAA) humano; el potenciador de troponina I (TN I); el potenciador del factor de crecimiento derivado de plaquetas; el potenciador de la distrofia muscular de Duchenne; el potenciador de poliovirus de SV40; el potenciador retroviral; el potenciador del virus del papiloma; el potenciador del virus de la hepatitis B; el potenciador de la inmunodeficiencia humana; el potenciador de citomegalovirus; y el potenciador del virus de la leucemia del mono gibbon.

Las secuencias de promotores/potenciadores inducibles ejemplares y de sus inductores se listan a continuación.

TABLA 3. Promotores/potenciadores inducibles útiles	
Elemento	Inductor
MTII	éster de forbol (TFA) de metales pesados
MMTV (virus del tumor mamario de ratón)	Glucocorticoides
Interferón- β	poli(r)Xpoli(rc)
Adenovirus 5 E2	Ela
c-jun	Éster de forbol (TPA), H ₂ O ₂
Colagenasa	Éster de forbol (TPA)
Estromelisin	Éster de forbol (TPA), IL-1
SV40	Éster de forbol (TFA)
Gen MX murino	Interferón, Virus de la enfermedad de Newcastle
Gen GRP78	A23187

TABLA 3. Promotores/potenciadores inducibles útiles	
Elemento	Inductor
α -2-macroglobulina	IL-6
Vimentina	Suero
Gen H-2kB de CMH de clase I	Interferón
HSP70	Ela Antígeno T grande de SV40
Proliferina	Éster de forbol (TPA)
Factor de necrosis tumoral	FMA
Gen α estimulante de la hormona tiroidea	Hormona tiroidea

Pueden obtenerse secuencias reguladoras de eucariotas adicionales a partir de la base de datos de promotores de eucariota EPDB, y también pueden usarse para dirigir la expresión de un polinucleótido.

- 5 En determinadas realizaciones de la invención, puede identificarse la administración de un vector en una célula *in vitro* o *in vivo* incluyendo un marcador de selección en la construcción de expresión, tal como se describe anteriormente en el presente documento. El marcador podría dar como resultado un cambio identificable en la célula modificada que permita la fácil identificación de la expresión. Normalmente, la inclusión de un marcador de selección de fármaco ayuda en el cierre y la selección de transformantes. Los genes que pueden usarse como marcadores de selección en células eucariotas se conocen en la técnica e incluyen, como ejemplos de marcadores de selección dominante, el gen de aminoglicósido 3' fosfotransferasa bacteriano (también citado como el gen neo) que confiere resistencia al fármaco G418 en células de mamífero, el gen de higromicina G fosfotransferasa (hyg) bacteriano que confiere resistencia al antibiótico higromicina y el gen de xantina-guanina fosforribosil transferasa bacteriano (también citado como el gen gpt) que confiere la capacidad de crecer en presencia de ácido micofenólico. Otros marcadores de selección no son dominantes en tanto que su uso debe ser conjunto con una línea celular que carezca de la actividad enzimática relevante. Los ejemplos de marcadores de selección no dominantes incluyen el gen de timidina cinasa (tk) que se usa en conjunción con líneas celulares tk⁻, el gen CAD que se usa en conjunción con células deficientes para CAD y el gen de hipoxantina-guanina fosforribosil transferasa (hprt) que se usa en conjunción con líneas celulares hprt⁻. Se proporciona una revisión del uso de marcadores de selección en líneas celulares de mamífero en Sambrook, J. *et al.*, 1989, *Molecular Cloning: A Laboratory Manual*, 2ª ed., (Cold Spring Harbor Laboratory Press, Nueva York págs.16.9-16.15.

25 Como alternativa, pueden emplearse genes que codifican enzimas, tales como timidina cinasa (tk) del virus del herpes simple (eucariotas) o cloranfenicol acetiltransferasa (CAT) (procariotas) para proporcionar marcadores de selección. También pueden emplearse marcadores inmunológicos. No se cree que el tipo exacto de marcador de selección sea importante, siempre que sea capaz de expresarse simultáneamente con el polinucleótido que codifica un polipéptido de interés. Los ejemplos adicionales de marcadores de selección se conocen bien por los expertos en la materia.

30 En los casos donde se emplea una inserción de ADNc para expresar un polipéptido de la familia de SPARC de la invención, típicamente se deseará incluir una señal de poliadenilación para efectuar la poliadenilación adecuada del transcrito de polinucleótido. No se cree que la naturaleza de la señal de poliadenilación sea crucial para la puesta en práctica exitosa de la invención, y puede emplearse cualquiera de estas secuencias. Estos elementos pueden servir para potenciar los niveles de mensaje y para minimizar la lectura desde el casete de expresión hacia otras secuencias.

Producción de células recombinantes: Introducción de un polinucleótido de la familia de SPARC en una célula para expresión y/o sensibilización

40 Tal como se ha descrito anteriormente, puede introducirse un polinucleótido de la familia de SPARC de la invención en una célula para expresar el polipéptido de la familia de SPARC para su purificación o para expresar y lograr el efecto de sensibilización de la invención de acuerdo con métodos bien conocidos en la técnica, por ejemplo, en *Current Protocols in Molecular Biology*, Ausubel, F.M. *et al.* (eds.), 1989, Greene Publishing Associates, Sección 9.1 y en *Molecular Cloning: A Laboratory Manual*, 2ª edición, Sambrook *et al.*, 1989, Cold Spring Harbor Laboratory Press.

50 En el presente documento se contemplan varios métodos no virales para la transferencia de vectores en células de mamífero cultivadas. Estos incluyen precipitación de fosfato de calcio (Graham, *et al.*, 1973; Chen, *et al.*, 1987; Rippe, *et al.*, 1990) DEAE-dextrano (Gopal, 1985), electroporación (Tur-Kaspa *et al.*, 1986; Potter *et al.*, 1984), microinyección directa (Harland, *et al.*, 1985), liposomas cargados de ADN (Nicolau, *et al.*, 1982; Fraley *et al.*, 1979) y complejos de ADN lipofectamina, sonicación celular (Fechheimer *et al.*, 1987), bombardeo génico usando microproyectiles de alta velocidad (Yang *et al.*, 1990), y transfección mediada por receptores (Wu, *et al.*, 1987; Wu, *et al.*, 1988). Algunas de estas técnicas pueden ser adaptarse satisfactoriamente para su uso *in vivo* o *ex vivo*.

Una vez que se ha administrado el vector a la célula, puede colocarse y expresarse el polinucleótido que codifica un polipéptido de la familia de SPARC en sitios diferentes. En determinadas realizaciones, el polinucleótido que codifica un polipéptido de la familia de SPARC puede integrarse de manera estable en el genoma de la célula. Esta integración puede ser mediante recombinación de homólogos (reemplazo génico) o puede integrarse en una ubicación al azar, no específica (aumento génico), véase Holmes-Son *et al.*, 2001, Adv. Genet. 43: 33-69. En otras realizaciones más, el polinucleótido que codifica un polipéptido de la familia de SPARC puede mantenerse establemente en la célula como un segmento de ADN episómico separado. Dichos segmentos de polinucleótido o "episomas" codifican secuencias suficientes para permitir el mantenimiento y la replicación independientemente de o en sincronía con el ciclo celular del sujeto. Se sabe bien cómo se administra la construcción de expresión a la célula y en qué parte de la célula permanece el polinucleótido, y esto depende del tipo de construcción de expresión empleada.

Las líneas celulares procedentes de especies de mamíferos que pueden ser adecuadas para la transfección e infección de un polinucleótido de la familia de SPARC y para la expresión y purificación de un polipéptido recombinante, pueden estar disponibles comercialmente. Estas líneas celulares incluyen, pero sin limitación, CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C1271 (ATCC CRL 1616), BS-C-1 (ATCC CCL 26), MRC-5 (ATCC CCL 171), células L, HEK-293 (ATCC CRL1573), NSO (ECACC85110503) y HT1080.

Los cultivos celulares pueden prepararse de distintos modos para la transferencia génica *in vitro*. Para que las células se mantengan viables *in vitro* y en contacto con la construcción de expresión, es necesario asegurar que las células mantengan contacto con la proporción correcta de oxígeno y dióxido de carbono y nutrientes pero que estén protegidos de la contaminación microbiana.

La transferencia de la construcción puede llevarse a cabo mediante cualquiera de los métodos conocidos en la técnica y tal como se describe más adelante en el presente documento. Algunos métodos pueden ser particularmente aplicables para la transferencia *in vitro* pero también pueden aplicarse para su uso *in vivo*.

Transfección mediada por CaPO₄

Puede introducirse en células un polinucleótido que codifica un polipéptido de la familia SPARC formando un precipitado que contiene el polinucleótido y fosfato de calcio. Por ejemplo, puede mezclarse una solución salina tamponada con HEPES con una solución que contiene cloruro de calcio y polinucleótido para formar un precipitado y entonces se incuba el precipitado con células. Puede añadirse una etapa de choque con glicerol o dimetil sulfóxido para aumentar la cantidad de polinucleótido captado por ciertas células. La transfección mediada por CaPO₄ puede usarse para transfectar de manera estable (o transitoria) a células y solo es aplicable para la modificación de células *in vitro*. Los protocolos para la transfección mediada por CaPO₄ pueden encontrarse en Current Protocols in Molecular Biology, Ausubel, F.M. *et al.* (eds.), 1989, Greene Publishing Associates, Sección 9.1 y en Molecular Cloning: A Laboratory Manual, 2ª edición; Sambrook *et al.*, 1989, (Cold Spring Harbor Laboratory Press, Secciones 16.32-16.40 u otros manuales de laboratorio convencionales.

Dubensky *et al.* (1984) inyectó satisfactoriamente ADN de poliomavirus en forma de precipitados de CaPO₄ en el hígado y bazo de ratones adultos y neonatos demostrando replicación viral activa e infección aguda. Benvenisty y Neshif (1986) demostraron también que la inyección intraperitoneal directa de plásmidos precipitados con CaPO₄ daba como resultado la expresión de los genes transfectados. Por lo tanto, el polinucleótido que codifica al polipéptido de la familia de SPARC también puede transferirse de un modo similar *in vivo* para expresar un polipéptido de la familia de SPARC deseado, tal como se ha descrito anteriormente.

Transfección mediada por DEAE-Dextrano

Puede introducirse un polinucleótido que codifica un polipéptido de la familia de SPARC en células formando una mezcla del polinucleótido y DEAE-dextrano e incubando la mezcla con las células. Puede añadirse una etapa de choque con dimetilsulfóxido o cloroquina para aumentar la cantidad de captación de polinucleótido. Los protocolos para la transfección mediada por DEAE-dextrano pueden encontrarse en Current Protocols in Molecular Biology, Ausubel, F.M. *et al.* (eds.), 1989, Greene Publishing Associates, Sección 9.2 y en Molecular Cloning: A Laboratory Manual, 2ª edición, Sambrook *et al.*, 1989, (Cold Spring Harbor Laboratory Press, Secciones 16.41-16.46 u otros manuales de laboratorio convencionales.

Electroporación

También puede introducirse un polinucleótido que codifica un polipéptido de la familia de SPARC en células incubando las células y el polinucleótido juntos en un tampón adecuado y sometiendo a las células a un pulso eléctrico de alto voltaje. La eficacia con la que se introduce el polinucleótido en las células mediante electroporación está influenciada por la fuerza del campo aplicado, la duración del pulso eléctrico, la temperatura, la conformación y la concentración del polinucleótido y la composición iónica del medio. La electroporación puede usarse para transfectar de manera estable (o transitoria) a una gran variedad de tipos celulares. Los protocolos para electroporar

células pueden encontrarse en Ausubel, F.M. *et al.* (eds.), *anteriormente citado*, Sección 9.3 y en Sambrook *et al.*, *anteriormente citado*, Secciones 16.54-16.55 u otros manuales de laboratorio convencionales.

Transfección mediada por liposomas ("Lipofección")

También puede introducirse un polinucleótido que codifica un polipéptido de la familia de SPARC en células mezclando el polinucleótido con una suspensión de liposomas que contiene lípidos catiónicos. El complejo polinucleótido/liposoma se incuba entonces con células. La transfección mediada por liposomas puede usarse para transfectar de manera estable (o transitoria) a células en cultivo *in vitro*. Los protocolos pueden encontrarse en Ausubel, F.M. *et al.* (eds.), *anteriormente citado*, Sección 9.4 y en otros manuales de laboratorio convencionales. Además, se ha logrado la administración de genes *in vivo* usando liposomas. Véase, por ejemplo, Nicolau *et al.*, 1987, *Meth. Enz.*, 4149:157-176; Wang, *et al.*, 1987, *Proc. Natl. Acad. Sci. USA*, 84:7851-7855; Brigham *et al.*, 1989, *Am. J. Med. Sci.*, 298:278; y Gould-Fogerite *et al.*, 1989, *Gene*, 84:429-438.

Inyección directa

Puede introducirse un polinucleótido que codifica un polipéptido de la familia de SPARC en células inyectando directamente el polinucleótido en las células. Para un cultivo de células *in vitro*, el polinucleótido puede introducirse mediante microinyección. Ya que se microinyecta a cada célula individualmente, esta estrategia es muy intensiva en cuanto al trabajo cuando se modifican grandes cantidades de células. Sin embargo, una situación en la que la microinyección es un método de elección es en la producción de animales transgénicos (discutido en más detalle más adelante). En esta situación, el polinucleótido se introduce de manera estable en un ovocito fertilizado que después se deja que se desarrolle en un animal. El animal resultante contiene células que portan el polinucleótido introducido en el ovocito. La inyección directa puede usarse para introducir el polinucleótido que codifica el polipéptido de la familia de SPARC en células *in vivo* (véase, por ejemplo, Acsadi *et al.*, 1991, *Nature*, 332: 815-818; Wolff *et al.*, 1990, *Science*, 247:1465-1468). Un aparato de administración (por ejemplo, una "pistola génica") puede usarse para inyectar ADN en células *in vivo*. Dicho aparato está disponible comercialmente (por ejemplo, a través de BioRad).

Captación de ADN mediada por receptor

Un polinucleótido que codifica el polipéptido de la familia de SPARC también puede introducirse en células formando un complejo del polinucleótido con un catión, tal como polilisina, que se acopla a un ligando para un receptor de la superficie celular (véase, por ejemplo Wu, *et al.*, 1988, *J. Biol. Chem.*, 263:14621; Wilson *et al.*, 1992, *J. Biol. Chem.*, 267:963-967; y la Patente de los Estados Unidos n.º 5.166.320). La unión del complejo polinucleótido-ligando al receptor facilita la captación del polinucleótido mediante endocitosis mediada por receptor. Los receptores a los que se dirigen el complejo polinucleótido-ligando incluyen el receptor de transferrina y el receptor de asialoglucoproteína. Un complejo de polinucleótido-ligando unido a cápsidas de adenovirus que altera a los endosomas de manera natural, de este modo liberando material en el citoplasma puede usarse para evitar la degradación del complejo por los lisosomas intracelulares (véase, por ejemplo, Curiel *et al.*, 1991, *Proc. Natl. Acad. Sci. USA*, 88:8850; Cristiano *et al.*, 1993, *Proc. Natl. Acad. Sci. USA*, 90:2122-2126). La captación de polinucleótido mediada por receptor puede usarse para introducir el polinucleótido que codifica el polipéptido de la familia de SPARC en células ya sea *in vitro* o *in vivo* y, además, tiene la característica añadida de que el polinucleótido puede dirigirse de manera selectiva a un tipo celular particular mediante el uso de un ligando que se une a un receptor expresado de manera selectiva en una célula diana de interés.

Transferencia génica mediada por virus

Otra estrategia para introducir el polinucleótido que codifica el polipéptido de la familia de SPARC en células es mediante el uso de un vector viral que contiene el polinucleótido que codifica el polipéptido de la familia de SPARC. La infección de células con un vector viral tiene la ventaja de que una gran proporción de células reciben el polinucleótido, que puede obviar la necesidad de selección de células que hayan recibido el polinucleótido. Además, las moléculas codificadas dentro del vector viral, por ejemplo, mediante un ADNc contenido en el vector viral, se expresan de manera eficaz en células que hayan captado el polinucleótido del vector viral y pueden usarse sistemas de vector viral bien *in vitro* o *in vivo*.

Los vectores virales no replicantes pueden producirse en líneas celulares empaquetadoras que producen partículas virales que son infecciosas pero defectuosas para replicación, haciendo que sean vectores útiles para la introducción de polinucleótidos en una célula que carezca de la información genética complementaria que permita la encapsidación (Mann *et al.*, 1983, *Cell*, 33:153; Miller y Buttimore, *Mol. Cell. Biol.*, 1986, 6:2895 (PA317, ATCC CRL9078). Se prefieren líneas celulares empaquetadoras que contengan genes empaquetadores anfotróficos capaces de transformar células de origen humano y de otras especies.

Retrovirus

Los retrovirus son un grupo de virus de ARN monocatenario caracterizados por una capacidad para convertir su ARN en ADN bicatenario en células infectadas mediante un proceso de retrotranscripción (Coffin, 1990, en Fields *et al.*, Ceds, Virology, Raven Press, Nueva York, págs. 1437-1500). El ADN resultante se integra entonces de manera estable en los cromosomas celulares en forma de provirus y dirige la síntesis de las proteínas virales. La integración da como resultado la retención de las secuencias génicas virales en la célula receptora y sus descendientes. El genoma retroviral contiene tres genes, *gag*, *pol*, y *env* que codifican las proteínas de la cápsida, la enzima polimerasa, y los componentes de la envuelta, respectivamente. Una secuencia hallada cadena arriba del gen *gag*, funciona como señal para empaquetar el genoma en viriones. Dos secuencias de repetición terminal larga (LTR) están presentes en los extremos 5' y 3' del genoma viral. Estas contienen secuencias promotoras y potenciadoras fuertes y también son necesarias para la integración en el genoma de la célula objeto (Coffin, *anteriormente citado*).

Los retrovirus defectuosos están bien caracterizados para su uso en la transferencia génica con fines de terapia génica (para una revisión, véase Miller, 1990, Blood 76:271).

Los protocolos para producir retrovirus recombinantes y para infectar células *in vitro* o *in vivo* con dichos virus pueden encontrarse en Current Protocols in Molecular Biology, Ausubel, F.M. *et al.* (eds.), 1989, Greene Publishing Associates, Secciones 9.10-9.14 y otros manuales de laboratorio convencionales. Los ejemplos de retrovirus adecuados incluyen pLJ, pZIP, pWE y pEM que se conocen bien por los expertos en la materia. Los ejemplos de líneas empaquetadoras de virus incluyen Crip, Cre, 2 y Am. Los retrovirus se han usado para introducir una diversidad de genes en muchos tipos celulares diferentes, incluyendo células epiteliales, células endoteliales linfocitos, mioblastos, hepatocitos, células de médula ósea, *in vitro* y/o *in vivo* (véase, por ejemplo, Eglitis, *et al.*, 1985, Science, 230:1395-1398; Danos, *et al.*, 1988, Proc. Natl. Acad. Sci. USA, 85:6460-6464; Wilson *et al.*, 1988, Proc. Natl. Acad. Sci. USA, 85:3014-3018; Armentano *et al.*, 1990, Proc. Natl. Acad. Sci. USA, 87:6141-6145; Huber *et al.*, 1991, Proc. Natl. Acad. Sci. USA, 88:8039-8043; Ferry *et al.*, 1991, Proc. Natl. Acad. Sci. USA, 88:8377-8381; Chowdhury *et al.*, 1991, Science, 254:1802-1805; van Beusechem *et al.*, 1992, Proc. Natl. Acad. Sci. USA, 89:7640-7644; Kay *et al.*, 1992, Human Gene Therapy, 3:641-647; Dai *et al.*, 1992, Proc. Natl. Acad. Sci. USA, 89:10892-10895; Hwu *et al.*, 1993, J. Immunol., 150:4104-4115; la Patente de Estados Unidos n.º 4.868.116; la Patente de Estados Unidos n.º 4.980.286; la solicitud PCT WO 89/07136; la solicitud PCT WO 89/02468; la solicitud PCT WO 89/05345; y la solicitud PCT WO 92/07573). Los vectores retrovirales necesitan de la división de la célula diana para que el genoma retroviral (y el polinucleótido exógeno insertado en este) se integre en el genoma objeto para introducir de manera estable el polinucleótido en la célula. Por lo tanto, puede ser necesario estimular la replicación de la célula diana.

Adenovirus

El conocimiento de la organización genética de adenovirus, un virus de ADN bicatenario lineal de 36 kB, permite la sustitución de un gran trozo de ADN adenoviral con secuencias exógenas de hasta 7 kB (Grunhaus, *et al.*, 1992, Seminar in Virology, 3:237-252). A diferencia de los retrovirus, la infección de ADN adenoviral en células diana no da como resultado la integración cromosómica ya que el ADN adenoviral puede replicarse de manera episómica sin una genotoxicidad potencial. También, los adenovirus son estructuralmente estables, y no se ha detectado reorganización del genoma después de una amplificación extensiva. Los adenovirus pueden infectar virtualmente a todas las células endoteliales, independientemente de su estado del ciclo celular.

Los adenovirus son particularmente adecuados para su uso como vector de transferencia génica debido a su genoma de tamaño medio, facilidad de manipulación, título elevado, amplio intervalo de células diana, y alta infectividad. Ambos extremos del genoma viral contienen repeticiones internas invertidas (ITR) de 100-200 pares de bases (pb), que son elementos *cis* necesarios para la replicación y empaquetamiento del ADN viral. Las regiones temprana (E) y tardía (L) del genoma contienen diferentes unidades de transcripción que se dividen mediante la aparición de la replicación del ADN viral. La región E1 (E1A y B1B) codifica proteínas responsables de la regulación de la transcripción del genoma viral y unos pocos genes celulares. La expresión de la región E2 (E2A y E2B) da como resultado la síntesis de las proteínas para la replicación del ADN viral. Estas proteínas están implicadas en la replicación del ADN, la expresión génica tardía, y el apagamiento de la célula hospedadora (Renan, 1990). Los productos de los genes tardíos, incluyendo la mayoría de las proteínas de la cápsida viral, se expresan únicamente después de un procesamiento significativo de un solo transcrito primario emitido por el promotor tardío principal (MLP). El MLP es particularmente eficaz durante la fase tardía de la infección, y todos los ARNm emitidos a partir de este promotor poseen una secuencia de líder tripartito (TL) 5' que los convierte en los ARNm preferidos para la traducción.

El genoma de un adenovirus puede manipularse de tal forma que codifica y expresa un producto génico de interés pero se inactiva en cuanto a su capacidad para replicarse en un ciclo vital viral lítico normal. Véase, por ejemplo, Berkner, *et al.*, 1988, BioTechniques, 6:616; Rosenfeld, *et al.*, 1991, Science, 252:431-434; y Rosenfeld, *et al.*, 1992, Cell, 68:143-155. Los vectores adenovirales adecuados procedentes de la cepa de adenovirus Ad de tipo 5 dl324 u otras cepas de adenovirus (por ejemplo, Ad2, Ad3, Ad7, etc.) se conocen bien por los expertos en la materia. Los adenovirus recombinantes son ventajosos en tanto que no requieren de la división celular para ser vehículos de

administración génica eficaces y pueden usarse para infectar a una gran variedad de tipos celulares, incluyendo el epitelio de las vías respiratorias (Rosenfeld, *et al.*, 1992, citado anteriormente), células endoteliales (Lemarchand, *et al.*, 1992, Proc. Natl. Acad. Sci. USA, 89:6482-6486), hepatocitos (Herz, *et al.*, 1993, Proc. Natl. Acad. Sci. USA, 90:2812-2816) y células musculares (Quantin, *et al.*, 1992, Proc. Natl. Acad. Sci. USA, 89:2581-2584). Además, el polinucleótido adenoviral introducido (y el ADN exógeno contenido en el mismo) no se integra en el genoma de una célula diana, sino que se mantiene episómico, evitando de este modo los problemas potenciales que pueden suceder como resultado de la mutagénesis de inserción en situaciones en las que el polinucleótido introducido se integra en el genoma diana (por ejemplo, ADN retroviral). Además, la capacidad portadora del genoma adenoviral para ADN exógeno es grande (de hasta 8 kilobases) en relación a otros vectores de administración génica (Berkner, *et al.* citado anteriormente; Haj-Ahmand, *et al.*, 1986, J. Virol., 57:267). La mayoría de los vectores adenovirales defectuosos para replicación actualmente en uso tienen eliminaciones de la totalidad o partes de los genes virales E1 y E3 pero mantienen un máximo del 80 % del material genético adenoviral.

Los adenovirus recombinantes pueden generarse mediante métodos conocidos en la técnica, por ejemplo, tal como se describe en la patente de Estados Unidos n.º 6.194.191.

La generación y propagación de los vectores adenovirales, que son deficientes para replicación, depende de una línea celular colaboradora única, denominada 293, que se transformó a partir de células de riñón embrionario humano mediante fragmentos de ADN de Ad5 y que expresa de manera constitutiva proteínas E1 (Graham, *et al.*, 1977). Ya que la región E3 es dispensable del genoma del adenovirus (Jones, *et al.*, 1978), los vectores de adenovirus actuales, con la ayuda de las células 293, portan ADN exógeno en E1, E3 o en ambas regiones (Graham, *et al.*, 1991). En la naturaleza, los adenovirus pueden empaquetar aproximadamente el 105% del genoma de tipo silvestre (Ghosh-Choudhury, *et al.*, 1987), proporcionando capacidad para aproximadamente 2 kB de ADN adicionales. En combinación con el ADN de aproximadamente 5,5 kB que puede reemplazarse en las regiones E1 y E3, la capacidad máxima para los vectores de adenovirus actuales es menor de 7,5 kB, o de aproximadamente el 15% de la longitud total del vector. Más del 80% del genoma viral de los adenovirus permanece en el amazón del vector y es la fuente de la citotoxicidad transmitida por vectores. También, la deficiencia en replicación del virus al que se le ha eliminado E1 es incompleta. Por ejemplo, se ha observado una filtración de la expresión génica viral con los vectores de adenovirus actualmente disponibles a altas multiplicidades de infección (Mulligan, 1993).

Las líneas celulares colaboradoras pueden proceder de células humanas, tales como células de riñón embrionario humano, células musculares, células hematopoyéticas u otras células mesenquimales o epiteliales embrionarias humanas. Como alternativa, las células colaboradoras pueden proceder de células de otras especies de mamífero que sean permisivas para adenovirus humanos. Dichas células incluyen, por ejemplo, células Vero u otras células mesenquimales o epiteliales embrionarias de mono. Como se afirma anteriormente, la línea celular colaboradora preferida es 293.

Aparte del requisito de que el vector de adenovirus sea defectuoso para replicación, o al menos condicionalmente defectuoso, no se cree que la naturaleza del vector adenoviral sea crucial para la puesta en práctica exitosa de la invención. El adenovirus puede ser de cualquiera de los 42 serotipos distintos conocidos o de los subgrupos A-F. Los adenovirus de tipo 5 del subgrupo C es el material de partida preferido para obtener el vector de adenovirus defectuoso para replicación condicional para su uso en el método de la presente invención. Esto se debe a que el adenovirus de tipo 5 es un adenovirus humano del cual se conoce una gran cantidad de información bioquímica y genética, y se ha usado históricamente para la mayoría de construcciones que emplean adenovirus como vector.

Como se afirma anteriormente, el vector típico, tal como se enseña en el presente documento es defectuoso para replicación y no tendrá una región E1 de adenovirus. Por lo tanto, será más conveniente introducir el polinucleótido que codifica un polipéptido de interés en la posición de la que se ha eliminado las secuencias codificantes de E1. Sin embargo, la posición de la inserción de la región codificante de un polinucleótido seleccionado dentro de las secuencias del adenovirus no es crítica.

Los adenovirus son fáciles de cultivar y manipular y muestran un gran intervalo de dianas *in vitro* e *in vivo*. Este grupo de virus puede obtenerse con títulos elevados, por ejemplo, 10^9 - 10^{11} unidades formadoras de placas (UFP)/ml, y son altamente infecciosos. El ciclo vital de los adenovirus no requiere de la integración en el genoma del sujeto.

Los vectores de adenovirus se han usado en la expresión de genes de eucariotas (Levero, *et al.*, 1991, Gene, 101:195-202; Gomez-Foix, *et al.*, 1992, J. Biol. Chem., 267:25129-25134) y en el desarrollo de vacunas (Grunhaus, *et al.*, 1992, Seminar in Virology, 3:237-252; Graham, *et al.*, 1992, Biotechnology, 20:363-390). Los estudios en animales sugirieron que pueden usarse adenovirus recombinantes para terapia génica (Stratford-Perricaudet, *et al.*, 1991, en: Human Gene Transfer, O. Cohen-Haguener, Ceds), John Libbey Eurotext, Francia; Stratford-Perricaudet, *et al.*, 1990, Hum. Gene Ther., 1:241-256; Rich, *et al.*, 1993, Nature, 361:647-650). Los experimentos de administración de adenovirus recombinantes a diferentes tejidos incluyen instilación traqueal (Rosenfeld, *et al.*, 1991, Science, 252:431-434; Rosenfeld, *et al.*, 1992, Cell, 68:143-155), inyección muscular (Ragot, *et al.*, 1993, Nature, 361:647-650), inyección intravenosa periférica (Herz, *et al.*, 1993, Proc. Natl. Acad. Sci. USA 90:2812-2816), e inoculación estereotáctica en el cerebro (Le Gal La Salle *et al.*, 1993, Science, 259:988-990).

Otros vectores virales como vectores de expresión

Pueden emplearse otros vectores virales como construcciones de expresión en la presente invención. Pueden usarse vectores procedentes de virus, tales como virus *vaccinia* (Ridgeway, 1988, en: Rodríguez R L, Denhardt D T, ed. *Vectors: A Survey of Molecular Cloning Vectors and Their Uses*. Stoneham: Butterworth, págs. 467-492; Baichwal, *et al.*, 1986 en: Kuchelapati R, ed. *Gene Transfer*. Nueva York: Plenum Press, págs. 117-148; Coupar, *et al.*, 1988, *Gene*, 68:1-10), virus adenoasociados (AAV) (Baichwal, *et al.*, 1986, *anteriormente citado*; Hemonot, *et al.*, 1984, *Proc. Natl. Acad. Sci. USA*, 81:6466-6470) y herpesvirus. Estos ofrecen varias características atractivas para diversas células de mamífero (Friedmann, 1989, *Science*, 244:1275-1281; Baichwal, *et al.*, 1986, *anteriormente citado*; Coupar, *et al.*, 1988, *anteriormente citado*; Horwich, *et al.*, 1990, *J. Virol.*, 64:642-650).

Virus adenoasociados: Los virus adenoasociados (AAV) son virus defectuosos de origen natural que requieren de otro virus, tal como un adenovirus o un herpesvirus, como virus auxiliar para la replicación eficaz y un ciclo vital productivo. (Para una revisión véase Muzyczka *et al.*, 1992, *Curr. Topics in Micro. and Immunol.*, 158:97-129). También es uno de los pocos virus que puede integrar su ADN en células que no están en división, y muestra una alta frecuencia de integración estable (véase, por ejemplo, Flotte *et al.*, 1992, *Am. J. Respir. Cell. Mol. Biol.*, 7:349-356; Samulski *et al.*, 1989, *J. Virol.*, 63:3822-3828; y McLaughlin *et al.*, 1989, *J. Virol.*, 62:1963-1973). Pueden empaquetarse y pueden integrarse vectores que contienen tan pocos como 300 pares de bases de AAV. El espacio para polinucleótido exógeno está limitado a aproximadamente 4,5 kb. Un vector de AAV, tal como el descrito en Tratschin *et al.*, 1985, *Mol. Cell. Biol.*, 5:3251-3260 puede usarse para introducir polinucleótidos en células. Se ha introducido una diversidad de polinucleótidos en diferentes tipos celulares usando vectores de AAV (véase, por ejemplo, Hemonot, *et al.*, 1984, *Proc. Natl. Acad. Sci. USA*, 81:6466-6470; Tratschin, *et al.*, 1985, *Mol. Cell. Biol.*, 4:2072-2081; Wondisford, *et al.*, 1988, *Mol. Endocrinol.*, 2:32-39; Tratschin, *et al.*, 1984, *J. Virol.*, 51:611-619; y Flotte, *et al.*, 1993, *J. Biol. Chem.*, 268:3781-3790).

Después de la transferencia de un polinucleótido que codifica el polipéptido de la familia de SPARC en células, las células pueden seleccionarse y usarse para sensibilizar para el tratamiento de acuerdo con la presente invención. La eficacia de un sistema y método de vector de expresión particular para introducir un polinucleótido en una célula puede evaluarse mediante estrategias convencionales usadas de manera rutinaria en la técnica. Por ejemplo, puede detectarse el polinucleótido introducido en una célula mediante una técnica de hibridación de filtro (por ejemplo, transferencia de Southern) y puede detectarse el ARN producido mediante la transcripción del polinucleótido introducido, por ejemplo, mediante transferencia de Northern, protección de RNasa o reacción en cadena de la polimerasa transcriptasa inversa (RT-PCR). El producto génico puede detectarse mediante un ensayo adecuado, por ejemplo, mediante la detección inmunológica de una proteína producida, tal como con un anticuerpo específico, o mediante un ensayo funcional para detectar una actividad funcional del producto génico, tal como un ensayo enzimático. Como alternativa, puede optimizarse en primer lugar un sistema de expresión usando un gen indicador unido a los elementos regulatorios y usarse el vector tal como se describe anteriormente en el presente documento. El gen indicador codifica un producto génico separado que es fácilmente detectable y, por lo tanto, puede usarse para evaluar la eficacia del sistema.

Ya que un polipéptido de la familia de SPARC es una proteína secretada, su nivel extracelular puede detectarse también para determinar su expresión usando métodos conocidos en la técnica, tales como inmunotransferencia o ELISA.

Otro modo de aumentar la expresión de un polinucleótido o polipéptido de SPARC en una célula es mediante la activación génica endógena, es decir, insertando un promotor fuerte antes de la secuencia del gen de la familia de SPARC natural en el genoma de la célula. La activación de genes endógenos es un método para introducir, mediante recombinación de homólogos con ADN genómico, secuencias de ADN (por ejemplo, promotores fuertes) que normalmente no están unidos funcionalmente al gen endógeno y (1) que, cuando se insertan en el genoma hospedador en o próximos al gen endógeno, sirven para alterar (por ejemplo, activar) la expresión del gen endógeno, y además (2) permiten la selección de células en las que el gen endógeno activado se amplifica. La expresión de proteínas mediante activación de genes endógenos se conoce bien en la técnica y se divulga, por ejemplo, en las Patentes de Estados Unidos n.º 5.733.761, 5.641.670, y 5.733.746, y en las publicaciones internacionales de patente n.º WO 93/09222, WO 94/12650, WO 95/31560, WO 90/11354, WO 91/06667 y WO 91/09955.

Tal como se enseña en el presente documento, la expresión de un gen de la familia de SPARC endógeno se activa (por ejemplo, aumenta) insertando un promotor/operador de tetraciclina inducible por tetraciclina para controlar su expresión.

Terapia para el cáncer

El cáncer se trata típicamente mediante cirugía, quimioterapia o radioterapia. También se están desarrollando terapias biológicas, tales como la inmunoterapia y la terapia génica. Otras terapias incluyen terapia hipertérmica, terapia fotodinámica, etc. (véase la página de bienvenida del National Cancer Institute en www.nci.nih.gov).

Quimioterapia

La quimioterapia es el uso de fármacos anticáncer (citotóxicos) para destruir células cancerosas. Hay más de 50 fármacos de quimioterapia diferentes y algunos se administran solos, pero a menudo pueden combinarse varios fármacos (eso se conoce como quimioterapia combinada). Una lista de ejemplo de agentes quimioterapéuticos, tal como se describe en la página www.cancerbacup.org.uk/info/actinomycin.htm, incluye: actinomicina D, adriamicina, altretamina, asparaginasa, bleomicina, busulfán, capecitabina, carboplatino, camustina, clorambucilo, cisplatino, CPT-11, ciclofosfamida, citarabina, dacarbazina, daunorrubicina, doxorubicina, epirubicina, etopósido, fludarabina, fluorouracilo, gemcitabina, hidroxiurea, idarubicina, fosfamida, irinotecán, doxorubicina liposómica, lomustina, melfalano, mercaptopurina, metotrexato, mitomicina, mitoxantrona, oxaliplatino, procarbazona, esteroides, estreptozocina, taxol, taxotere, taxotere, el ensayo TACT, tamoxolomida, tioguanina, tiotepa, tomudex, topotecán, treosulfán, UFT (Uracilo-tegufur), vinblastina, vincristina, vindesina, vinorelbina.

Debido a que las células cancerosas pueden crecer y dividirse más rápidamente que las células normales, se producen muchos fármacos anticáncer para que eliminen a células en crecimiento. Pero algunas células normales sanas también se multiplican rápidamente, y la quimioterapia puede afectar también a estas células. Este daño a células normales provoca efectos secundarios. Las células normales de crecimiento rápido que pueden verse afectadas con mayor facilidad son las células sanguíneas que se forman en la médula ósea y las células del tracto digestivo (boca, estómago, intestino, esófago), aparato reproductor (órganos sexuales), y folículos pilosos. Algunos fármacos anticáncer pueden afectar a células de los órganos vitales, tales como el corazón, riñón, vejiga, pulmones, y sistema nervioso.

Los tipos de efectos secundarios que se tienen y la gravedad de estos dependen del tipo y dosis de la quimioterapia que se administre y de cómo reacciona el organismo. Los efectos secundarios de la quimioterapia incluyen fatiga, náuseas y vómitos, dolor, pérdida de cabello, anemia, problemas del sistema nervioso central, infección, problemas de coagulación sanguínea, problemas de boca, encías y garganta, diarrea, estreñimiento, efectos nerviosos y musculares, efectos en la piel y uñas, hipersensibilización a la radiación, efectos en el riñón y la vejiga, síntomas parecidos a la gripe, y retención de líquidos.

Radioterapia

Un tipo de radioterapia usada comúnmente implica fotones, "paquetes" de energía. Los rayos X fueron la primera forma de radiación fotónica para su uso en el tratamiento del cáncer. Dependiendo de la cantidad de energía que posean, los rayos pueden usarse para destruir células cancerosas en la superficie o en el interior del organismo. Cuanta más energía tenga el haz de rayos X, más profundamente podrán penetrar los rayos X en el tejido diana. Los aceleradores lineales y los betatrones son máquinas que producen rayos X con una energía cada vez mayor. El uso de máquinas para enfocar la radiación (tal como rayos X) en un sitio con cáncer se llama radioterapia de haz externo.

Los rayos gamma son otra forma de protones usada en la radioterapia. Los rayos gamma se producen espontáneamente a medida que determinados elementos (tales como radio, uranio, y cobalto 60) liberan radiación conforme se van descomponiendo, o desintegrando. Cada elemento se desintegra a una velocidad específica y emite energía en forma de rayos gamma y otras partículas. Los rayos X y gamma tienen el mismo efecto en las células cancerosas.

Otra técnica para administrar radiación a células cancerosas es la colocación de implantes radioactivos directamente en un tumor o cavidad corporal. Esto se denomina radioterapia interna. La braquiterapia, irradiación intersticial, y la irradiación intracavidad son tipos de radioterapia interna. En este tratamiento, la dosis de radiación se concentra en un área pequeña, y el paciente permanece en el hospital durante unos días. La radioterapia interna se usa frecuentemente para cánceres de lengua, de útero, y de cuello de útero.

Se están evaluando varias nuevas estrategias de radioterapia para determinar su eficacia para el tratamiento del cáncer. Una de dichas técnicas es la irradiación intraoperativa, en la que se dirige una gran dosis de irradiación externa sobre el tumor y el tejido circundante durante la cirugía.

Otra estrategia en investigación es el uso de radioterapia de haz de partículas. Este tipo de terapia difiere de la radioterapia fotónica en tanto que implica el uso de partículas subatómicas de movimiento rápido para tratar cánceres localizados. Se necesita una máquina sofisticada para producir y acelerar las partículas necesarias para este procedimiento. Algunas partículas (neutrones, piones, e iones pesados) depositan más energía a su paso a través del tejido que los rayos X o los rayos gamma, provocando mayores daños en las células sobre las que impactan. Este tipo de radiación se cita a menudo como radiación de alta transferencia lineal de energía (alta TLE).

Se están estudiando dos tipos de fármacos de investigación respecto de su efecto en células que se someten a radiación. Los radiosensibilizantes hacen que las células tumorales tengan más probabilidad de sufrir daños, y los radioprotectores protegen a los tejidos normales de los efectos de la radiación. La hipertermia, el uso de calor, también se está estudiando respecto de su eficacia para sensibilizar el tejido a la radiación.

Pueden usarse implantes de semillas radiactivas como la única modalidad de tratamiento para el adenocarcinoma de próstata para los pacientes adecuados con la enfermedad en estadio temprano. Las dos fuentes más comunes son yodo 125 y paladio 103, sin datos clínicos destacables acerca de que uno sea mejor que el otro. El implante de semillas radiactivas puede personalizarse individualmente para la próstata de un paciente para maximizar la dosis a la glándula a la vez que se minimiza la dosis a las estructuras normales circundantes. La braquiterapia de próstata ofrece el mayor nivel de radioterapia conformacional para el adenocarcinoma de próstata. El equipo de braquiterapia de próstata en la Thomas Jefferson University tiene una gran experiencia en braquiterapia de próstata y ha presentado trabajos en foros nacionales e internacionales.

La braquiterapia de próstata o implante de semillas radiactivas es un método dependiente del operario altamente técnico que administra la energía de la radiación mediante la colocación de pequeñas semillas radiactivas directamente dentro de la próstata, administrando eficazmente el tratamiento de "dentro hacia fuera". Esto se efectúa en el quirófano con anestesia general, como un procedimiento de una sola vez. Estas semillas pueden administrar grandes dosis de radiación directamente al tumor, con poco daño al tejido sano normal alrededor de la próstata. Esto puede combinarse en algunos casos con radioterapia conformacional en 3 dimensiones.

Otras investigaciones en radioterapia recientes se han centrado en el uso de anticuerpos radiomarcados para administrar dosis de radiación directamente al sitio de cáncer (radioinmunoterapia). Los anticuerpos son proteínas altamente eficaces que se producen por el cuerpo en respuesta a la presencia de antígenos (sustancias reconocidas como exógenas por el sistema inmunitario). Algunas células tumorales contienen antígenos específicos que desencadenan la producción de anticuerpos específicos de tumor. Pueden producirse grandes cantidades de estos anticuerpos en el laboratorio y unirse a sustancias radiactivas (un proceso conocido como radiomarcado). Una vez que se inyectan en el organismo, los anticuerpos buscan activamente a las células cancerosas, que se destruyen por la acción eliminadora de células (citotóxica) de la radiación. Esta estrategia puede minimizar el riesgo de daño por radiación a las células sanas. El éxito de esta técnica dependerá tanto de la identificación de sustancias radiactivas adecuadas y la determinación de la dosis segura y eficaz de radiación que puede administrarse de este modo.

La radioterapia puede usarse sola o en combinación con la quimioterapia y la cirugía. Al igual que todas las formas de tratamiento para el cáncer, la radioterapia puede tener efectos secundarios. Los posibles efectos secundarios del tratamiento con radiación incluyen pérdida temporal o permanente del cabello en la zona a tratar, irritación de la piel, cambios en la pigmentación de la piel en la zona tratada, y cansancio. Otros efectos secundarios dependen en gran medida de la zona del organismo que se esté tratando.

Terapia de hipertermia

La hipertermia, un procedimiento en el que el tejido corporal se expone a altas temperaturas (de hasta 41,1°C), se está investigando para evaluar su eficacia en el tratamiento del cáncer. El calor puede ayudar a reducir los tumores dañando a las células o privándolas de sustancias que necesitan para vivir. La terapia hipotérmica puede ser local, regional, e hipertermia en todo el organismo, usando dispositivos de calentamiento internos y externos. La hipertermia se usa prácticamente siempre con otras formas de terapia (radioterapia, quimioterapia, y terapia biológica) para intentar aumentar su eficacia.

La hipertermia local se refiere al calor que se aplica a una zona muy pequeña, tal como un tumor. La zona puede calentarse de manera externa con ondas de alta frecuencia dirigidas a un tumor desde un dispositivo extracorpóreo. Para lograr el calentamiento interno, puede usarse una de varios tipos de sondas estériles, incluyendo alambres finos calentados o tubos huecos rellenos de agua caliente; antenas de microondas implantadas; y electrodos de radiofrecuencia.

En la hipertermia regional, se calienta un órgano o extremidad. Se colocan imanes y dispositivos que producen alta energía sobre la zona que se va a tratar. En otro enfoque, llamado perfusión, se extrae parte de la sangre de un paciente, se calienta, y después se bombea (perfundes) en la zona que se va a tratar de manera interna.

El calentamiento del organismo completo se usa para tratar el cáncer metastásico que se ha diseminado por todo el organismo. Puede lograrse usando mantas de agua caliente, cera caliente, bobinas de inducción (tales como aquellas en las mantas eléctricas), o cámaras térmicas (similares a grandes incubadores).

La hipertermia no provoca un aumento destacable en los efectos secundarios o complicaciones de la radiación. El calor aplicado directamente a la piel, sin embargo, puede provocar malestar o incluso un dolor local significativo en aproximadamente la mitad de los pacientes tratados. También puede provocar ampollas, que en general se curan rápidamente. De manera menos común, puede causar quemaduras.

Terapia fotodinámica

La terapia fotodinámica (también conocida como TFD, fotorradioterapia, fototerapia, o fotoquimioterapia) es un tratamiento de algunos tipos de cáncer. Se basa en el descubrimiento de que determinados agentes químicos, conocidos como agentes fotosensibilizantes, pueden eliminar organismos unicelulares cuando se exponen los

organismos a un tipo de luz particular. La TFD destruye a las células cancerosas mediante el uso de una luz láser de frecuencia fija en combinación con un agente fotosensibilizante.

En la TFD, el agente fotosensibilizante se inyecta en el torrente sanguíneo y se absorbe por las células en todo el organismo. El agente permanece en las células cancerosas durante más tiempo que en las células normales. Cuando las células cancerosas tratadas se exponen a la luz láser, el agente fotosensibilizante absorbe la luz y produce una forma activa de oxígeno que destruye a las células cancerosas tratadas. La exposición a la luz tiene que cronometrarse cuidadosamente de tal forma que tenga lugar cuando la mayoría del agente fotosensibilizante haya abandonado las células sanas pero siga presente en las células cancerosas.

La luz láser usada en la TFD puede dirigirse a través de fibra óptica (una hebra de vidrio muy fina). La fibra óptica se coloca próxima al cáncer para administrar la cantidad de luz adecuada. La fibra óptica puede dirigirse mediante un broncoscopio a los pulmones para el tratamiento del cáncer de pulmón o mediante un endoscopio al esófago para el tratamiento del cáncer de esófago.

Una ventaja de la TFD es que causa un daño mínimo al tejido sano. Sin embargo, debido a que la luz láser usada actualmente no puede atravesar más de aproximadamente 3 centímetros de tejido, la TFD se usa principalmente para tratar tumores sobre o justo debajo de la piel o en el revestimiento de los órganos internos.

La terapia hace que la piel y los ojos sean fotosensibles durante 6 semanas o más después del tratamiento. Se aconseja a los pacientes que eviten la luz solar directa y la luz interior brillante durante al menos 6 semanas. Si los pacientes tienen que salir al exterior, necesitan llevar ropa protectora, incluyendo gafas de sol. Otros efectos secundarios temporales de la TFD están relacionados con el tratamiento de zonas específicas y pueden inducir tos, problemas de deglución, dolor abdominal, y respiración dolorosa o falta de aliento.

En diciembre de 1995, la U.S. Food and Drug Administration (FDA) aprobó un agente fotosensibilizante denominado porfímero sódico, o Photofrin®, para aliviar los síntomas del cáncer de esófago que está causando una obstrucción y para el cáncer de esófago que no puede tratarse de manera satisfactoria solo con láseres. En enero de 1998, la FDA aprobó el porfímero sódico para el tratamiento del cáncer de pulmón no microcítico temprano en pacientes para los que no son adecuados los tratamientos habituales para el cáncer de pulmón. El National Cancer Institute y otras instituciones están patrocinando ensayos clínicos (estudios de investigación) para evaluar el uso de la terapia fotodinámica para varios tipos de cáncer, incluyendo cánceres de vejiga, de cerebro, de laringe, y de la cavidad oral.

Terapia láser

La terapia láser implica el uso de luz de alta intensidad para destruir células cancerosas. Esta técnica se usa a menudo para aliviar los síntomas del cáncer, tales como sangrados u obstrucciones, especialmente cuando el cáncer no puede curarse mediante otros tratamientos. También puede usarse para tratar el cáncer reduciendo o destruyendo tumores.

El término "láser" significa amplificación de luz mediante emisión de radiación estimulada. La luz convencional, tal como la de una bombilla, tiene muchas longitudes de onda y se disemina en todas las direcciones. La luz láser, por su parte, tiene una longitud de onda específica y se enfoca en un haz estrecho. Este tipo de luz de alta intensidad contiene mucha energía. Los láseres son muy potentes y pueden usarse para cortar acero o para tallar diamantes. Los láseres también pueden usarse para un trabajo quirúrgico muy preciso, tal como reparar una retina dañada en el ojo o cortar tejido (en lugar de un bisturí).

Aunque hay varios tipos de láser diferentes, solo tres tipos se usan ampliamente en medicina:

- Láser de dióxido de carbono (CO₂) - Este tipo de láser puede retirar finas capas de la superficie de la piel sin penetrar en capas más profundas. Esta técnica es particularmente útil para tratar tumores que no se han diseminado profundamente en la piel y para determinadas afecciones precancerosas. Como alternativa a la cirugía con bisturí tradicional, el láser de CO₂ también es capaz de cortar la piel. El láser se usa de este modo para extirpar cánceres de piel.

- Láser de neodimio:itrio-aluminio-granate (Nd:YAG) - La luz de este láser puede penetrar más profundamente en tejidos que la luz de otros tipos de láser, y puede provocar la coagulación de la sangre rápidamente. Puede llevarse a través de fibra óptica a partes menos accesibles del organismo. Este tipo de láser se usa en ocasiones para tratar cánceres de garganta.

- Láser de argón - Este láser solo puede atravesar capas superficiales de tejido y por lo tanto es útil en dermatología y en cirugía ocular. También se usa con tintes fotosensibles para tratar tumores en un procedimiento conocido como terapia fotodinámica (TFD).

Los láseres tienen varias ventajas frente a las herramientas quirúrgicas convencionales, incluyendo:

- Los láseres son más precisos que los bisturíes. Se protege al tejido próximo a una incisión, ya que hay poco contacto con la piel circundante u otros tejidos.

- El calor producido por los láseres esteriliza la zona quirúrgica, reduciendo de este modo el riesgo de infección.

- Puede necesitarse un tiempo de intervención menor debido a que la precisión del láser permite una incisión más pequeña.

- A menudo se reduce el tiempo de curación; ya que el láser sella los vasos sanguíneos, hay menos sangrado, inflamación, o formación de cicatrices.

- La cirugía láser puede ser menos complicada. Por ejemplo, con fibra óptica, la luz láser puede dirigirse a partes del organismo sin hacer una gran incisión.

- Pueden efectuarse más procedimientos de tipo ambulatorio.

También hay desventajas con la cirugía láser:

- Relativamente pocos cirujanos están entrenados en el uso del láser.

- El equipamiento del láser es caro y voluminoso en comparación con las herramientas quirúrgicas convencionales, tales como los bisturíes.

- Deben observarse precauciones de seguridad estrictas en la sala de operaciones. (Por ejemplo, el equipo quirúrgico y el paciente tienen que llevar protección ocular).

Los láseres pueden usarse de dos modos para tratar el cáncer: reduciendo o destruyendo el tumor con calor, o activando un agente químico, conocido como agente fotosensibilizante, que destruye a las células cancerosas. En la TFD, se retiene un agente fotosensibilizante en las células cancerosas y puede estimularse con luz para provocar una reacción que elimina a las células cancerosas.

Los láseres de CO₂ y de Nd:YAG se usan para reducir o destruir tumores. Pueden usarse con endoscopios, tubos que permiten a los médicos ver determinadas partes del organismo, tales como la vejiga. La luz de algunos láseres puede transmitirse a través de un endoscopio flexible equipado con fibra óptica. Esto permite a los médicos ver y trabajar en partes del organismo que no podrían alcanzarse de otro modo salvo por cirugía y por lo tanto permite apuntar el haz de luz láser de manera muy precisa. Los láseres también pueden usarse con microscopios de bajo aumento, proporcionando al doctor una visión clara del sitio que se está tratando. Usados con otros instrumentos, los sistemas láser pueden producir un área de corte tan pequeña como 200 micras de diámetro, menos que el ancho de un filamento muy fino.

Los láseres se usan para tratar muchos tipos de cáncer. La cirugía láser es un tratamiento estándar para determinados estadios de cánceres de la glotis (cuerdas vocales), de cuello de útero, de piel, de pulmón, vaginal, vulvar, y de pene.

Además de su uso para destruir el cáncer, la cirugía láser también se usa para ayudar a aliviar los síntomas causados por el cáncer (cuidados paliativos). Por ejemplo, los láseres pueden usarse para reducir o destruir un tumor que está bloqueando la tráquea de un paciente (vías aéreas), facilitándole la respiración. También se usa en ocasiones para paliar el cáncer colorrectal y anal.

La termoterapia intersticial inducida por láser (LITT) es uno de los desarrollos más recientes en la terapia para el cáncer. La LITT usa la misma idea que un tratamiento para el cáncer denominado hipertermia; el calor puede ayudar a reducir los tumores dañando a las células o privándolas de sustancias que necesitan para vivir. En este tratamiento, los láseres se dirigen a zonas intersticiales (zonas entre los órganos) en el organismo. La luz láser entonces eleva la temperatura del tumor, que daña o destruye a las células cancerosas.

Terapia génica

La terapia génica es una intervención médica experimental que implica modificar el material genético de células vivas para combatir enfermedades. La terapia génica se está estudiando en ensayos clínicos (trabajos de investigación con seres humanos) para muchos tipos de cáncer diferentes y para otras enfermedades.

Una de las metas de la terapia génica es suministrar células con copias sanas de genes ausentes o alterados. En lugar de administrar un fármaco al paciente, los médicos intentan corregir el problema alterando la configuración genética de algunas de las células del paciente. Los ejemplos de enfermedades que podrían tratarse de este modo incluyen la fibrosis quística y la hemofilia.

La terapia génica también se está estudiando como una forma para cambiar el funcionamiento de una célula; por ejemplo, estimulando a células del sistema inmunitario para que ataquen a células cancerosas.

5 En general, se administra un gen a la célula usando un "vector". Los tipos de vectores más comunes usados en la terapia génica son virus. Los virus usados como vectores en terapia génica están genéticamente inhabilitados; son incapaces de reproducirse por sí mismos. La mayoría de los ensayos clínicos de terapia génica se basan en retrovirus de ratón para administrar el gen deseado. Otros virus usados como vectores incluyen adenovirus, virus adeno asociados, poxvirus, y virus del herpes.

10 Una terapia génica puede efectuarse tanto *ex vivo* como *in vivo*. En la mayoría de ensayos clínicos de terapia génica *ex vivo*, se extraen células de la sangre o de la médula ósea del paciente y se cultivan en laboratorio. Las células se exponen al virus que porta el gen deseado. El virus entra en las células, y el gen deseado se convierte en parte del ADN de la célula. Las células crecen en el laboratorio y entonces se devuelven al paciente mediante inyección en una vena. En la terapia génica *in vivo*, se usan vectores o liposomas para administrar el gen deseado a las células
15 dentro del organismo del paciente.

Inmunoterapia

20 El cáncer puede desarrollarse cuando el sistema inmunitario decae o no funciona de manera adecuada. La inmunoterapia usa el sistema inmunitario del organismo, de forma directa o indirecta, para combatir el cáncer o para aminorar los efectos secundarios que pueden provocar determinados tratamientos para el cáncer. La inmunoterapia está diseñada para reparar, estimular, o potenciar las respuestas del sistema inmunitario.

25 Las células del sistema inmunitario incluyen a las siguientes: Los linfocitos son un tipo de glóbulo blanco de la sangre encontrado en la sangre y en muchas otras partes del organismo. Los tipos de linfocitos incluyen células B, células T, y células citotóxicas naturales. Las células B (linfocitos B) maduran en células plasmáticas que secretan anticuerpos (inmunoglobulinas), las proteínas que reconocen y se unen a sustancias exógenas conocidas como antígenos. Cada tipo de célula B produce un anticuerpo específico, que reconoce a un antígeno específico. Las células T (linfocitos T) atacan directamente a células infectadas, exógenas, o cancerosas. Las células T también
30 regulan la respuesta inmunitaria señalizando para otros defensores del sistema inmunitario. Las células T funcionan principalmente produciendo proteínas llamadas linfocinas. Las células citotóxicas naturales (células NK) producen sustancias químicas potentes que se unen y eliminan a cualquier invasor exógeno. Atacan sin tener que reconocer en primer lugar a un antígeno específico. Los monocitos son glóbulos blancos de la sangre que pueden tragarse y digerir a organismos y partículas microscópicas en un proceso conocido como fagocitosis. Los monocitos pueden
35 desplazarse al interior de tejidos y convertirse en macrófagos.

Las células del sistema inmunitario secretan dos tipos de proteínas: anticuerpos y citocinas. Los anticuerpos responden a los antígenos enganchándose, o uniéndose, a antígenos. Los anticuerpos específicos se emparejan con antígenos específicos, acoplándose entre sí del modo que una llave se acopla a una cerradura. Las citocinas
40 son sustancias producidas por algunas células del sistema inmunitario para comunicarse con otras células. Los tipos de citocinas incluyen linfocinas, interferones, interleucinas, y factores estimulantes de colonias. Las citocinas citotóxicas se liberan por un tipo de célula T llamado célula T citotóxica. Estas citocinas atacan a las células cancerosas directamente.

45 Los agentes inmunomoduladores inespecíficos son sustancias que estimulan o aumentan indirectamente al sistema inmunitario. A menudo, estos agentes se dirigen a células del sistema inmunitario clave y provocan respuestas secundarias, tales como producción aumentada de citocinas e inmunoglobulinas. Dos agentes inmunomoduladores inespecíficos usados en el tratamiento del cáncer son el bacilo de Calmette-Guerin (BCG) y levamisol. El BCG, que se ha usado ampliamente como vacuna para la tuberculosis, se usa en el tratamiento del cáncer de vejiga superficial
50 después de la cirugía. El BCG puede funcionar estimulando una respuesta inflamatoria, y posiblemente una inmunitaria. Se instila una solución de BCG en la vejiga y se retiene durante aproximadamente 2 horas antes de que se deje al paciente vaciar la vejiga por micción. Este tratamiento se efectúa normalmente una vez a la semana durante 6 semanas. El levamisol se usa junto con la quimioterapia de fluorouracilo (5-FU) en el tratamiento del cáncer de colon en estadio III (C de Dukes) después de la cirugía. El levamisol puede actuar para restaurar la
55 función inmunitaria deprimida.

Algunos anticuerpos, citocinas, y otras sustancias del sistema inmunitario pueden producirse en el laboratorio para su uso en el tratamiento del cáncer. Estas sustancias se denominan a menudo modificadores de la respuesta biológica (MRB). Alteran la interacción entre las defensas inmunitarias del organismo y las células cancerosas para
60 potenciar, dirigir, o restaurar la capacidad del cuerpo para combatir la enfermedad. Los MRB incluyen interferones, interleucinas, factores estimuladores de colonias, anticuerpos monoclonales, y vacunas. La inmunoterapia puede usarse para detener, controlar, o suprimir procesos que permiten el crecimiento del cáncer; hacer que las células cancerosas sean más reconocibles, y por lo tanto más susceptibles, a la destrucción por parte del sistema inmunitario; potenciar la potencia de la eliminación de las células del sistema inmunitario, tales como las células T,
65 las células NK, y los macrófagos; alterar los patrones de crecimiento del cáncer para promover el comportamiento como el de las células sanas; bloquear o revertir el proceso que convierte a una célula normal o a una célula

precancerosa en una célula cancerosa; potenciar la capacidad del organismo para reparar o sustituir a células normales dañadas o destruidas por otras formas de tratamiento para el cáncer, tales como la quimioterapia o la radiación; y prevenir que las células cancerosas se diseminen a otras partes del organismo.

- 5 Algunos MRB son una parte convencional del tratamiento para determinados tipos de cáncer, mientras que otros se están estudiando en ensayos clínicos. Los MRB se están usando solos o en combinación con otros. También se están usando con otros tratamientos, tales como radioterapia y quimioterapia.

10 Los interferones (IFN) son tipos de citocinas que se producen naturalmente en el organismo. Fueron las primeras citocinas producidas en el laboratorio para su uso como MRB. Hay tres tipos principales de interferones, interferón alfa, interferón beta, e interferón gamma; el interferón alfa es el tipo más ampliamente usado en el tratamiento para el cáncer. Los interferones pueden mejorar el modo en que el sistema inmunitario de un paciente actúa contra las células cancerosas. Además, los interferones pueden actuar directamente en las células cancerosas frenando su crecimiento o promoviendo su desarrollo en células con un comportamiento más normal. Algunos interferones también pueden estimular a las células NK, células T, y macrófagos, potenciando la función anticáncer del sistema inmunitario. La U.S. Food and Drug Administration (FDA) ha aprobado el uso de interferón alfa para el tratamiento de determinados tipos de cáncer, incluyendo tricoleucemia, melanoma, leucemia mieloide crónica, y sarcoma de Kaposi relacionado con el SIDA. Los estudios han demostrado que el interferón alfa también puede ser eficaz para tratar otros cánceres, tales como el cáncer de riñón metastásico y el linfoma no de Hodgkin.

20 Al igual que los interferones, las interleucinas (IL) son citocinas que aparecen de manera natural en el organismo y pueden producirse en laboratorio. Se han identificado muchas interleucinas; la interleucina-2 (IL-2 o aldesleucina) ha sido el tratamiento para el cáncer más ampliamente estudiado. La IL-2 estimula el crecimiento y la actividad de muchas células inmunitarias, tales como linfocitos, que pueden destruir a las células cancerosas. La FDA ha aprobado IL-2 para el tratamiento del cáncer de riñón metastásico y del melanoma metastásico.

30 Los factores estimulantes de colonias (CSF) (en ocasiones denominados factores de crecimiento hematopoyéticos) no afectan normalmente de manera directa a las células tumorales; en cambio, inducen a las células de la médula ósea para que se dividan y se desarrollen en glóbulos blancos sanguíneos, plaquetas, y glóbulos rojos sanguíneos. La médula ósea es crítica para el sistema inmunitario del organismo ya que es la fuente de todas las células sanguíneas. La estimulación con CSF del sistema inmunitario puede beneficiar a pacientes que se sometan a tratamiento para el cáncer. Debido a que los fármacos anticáncer pueden dañar la capacidad del organismo para producir glóbulos blancos sanguíneos, glóbulos rojos sanguíneos, y plaquetas, los pacientes que reciben fármacos anticáncer tienen un riesgo aumentado de desarrollar infecciones, de desarrollar anemia, y de tener sangrados más fácilmente. Al usar CSF para estimular la producción de células sanguíneas, los médicos pueden aumentar las dosis de los fármacos anticáncer sin aumentar el riesgo de infección o la necesidad de transfusión con derivados de la sangre. Los CSF son particularmente útiles cuando se combinan con quimioterapia a altas dosis. Algunos ejemplos de CSF y su uso en la terapia para el cáncer son los siguientes: G-CSF (filgrastim) y GM-CSF (sargramostim) pueden aumentar el número de glóbulos blancos de la sangre, de este modo reduciendo el riesgo de infección en pacientes que reciben quimioterapia. G-CSF y GM-CSF también pueden estimular la producción de células madre en la preparación de trasplantes de células madre o de médula ósea; la eritropoyetina puede aumentar el número de glóbulos rojos de la sangre y reducir la necesidad de transfusiones de glóbulos rojos en pacientes que reciban quimioterapia; y Oprevelkin puede reducir la necesidad de transfusiones de plaquetas en pacientes que reciban quimioterapia.

45 Los CSF se usan en ensayos clínicos para tratar algunos tipos de leucemia, cáncer colorrectal metastásico, melanoma, cáncer de pulmón, y otros tipos de cáncer.

50 Los anticuerpos monodonaes (MOAB) también se están evaluando en la terapia para el cáncer. Estos anticuerpos se producen por un solo tipo de célula y son específicos para un antígeno particular. Se están creando MOAB específicos para los antígenos encontrados en la superficie de las células cancerosas.

55 Los MOAB se producen inyectando células cancerosas humanas en ratones, de tal forma que sus sistemas inmunitarios producirán anticuerpos contra estas células cancerosas. Las células de ratón que producen los anticuerpos se extraen entonces y se fusionan con células cultivadas en laboratorio para crear células "híbridas" llamadas hibridomas. Los hibridomas pueden producir de manera indefinida grandes cantidades de estos anticuerpos puros, o MOAB. Los MOAB pueden usarse en el tratamiento del cáncer de una serie de formas: Los MOAB que reaccionan con tipos específicos de cáncer pueden potenciar la respuesta inmunitaria del paciente al cáncer. Los MOAB pueden programarse para que actúen contra factores de crecimiento celulares, de este modo interfiriendo con el crecimiento de las células cancerosas. Los MOAB pueden unirse a fármacos anticáncer, radioisótopos (sustancias radioactivas), otros MRB, u otras toxinas. Cuando los anticuerpos se anclan a células cancerosas, administran estos venenos directamente al tumor, ayudando a destruirlo. Los MOAB pueden ayudar a destruir células cancerosas en la médula ósea que se ha retirado de un paciente en la preparación de un trasplante de médula ósea. Los MOAB que portan radioisótopos también pueden ser útiles para diagnosticar determinados cánceres, tales como colorrectal, ovárico, y de próstata.

Rituxan® (rituximab) y Herceptin® (trastuzumab) son ejemplos de anticuerpos monoclonales que han sido aprobados por la FDA. Rituxan se usa para el tratamiento de linfoma no de Hodgkin de células B que ha recaído tras un periodo de mejora o que no ha respondido a la quimioterapia. Herceptin se usa para tratar el cáncer de mama metastásico en pacientes con tumores que producen cantidades excesivas de una proteína llamada HER-2. (Aproximadamente el 25 por ciento de los tumores de cáncer de mama producen cantidades excesivas de HER-2). Se están comenzando a probar los MOAB en ensayos clínicos para tratar linfomas, leucemias, cáncer colorrectal, cáncer de pulmón, tumores cerebrales, cáncer de próstata, y otros tipos de cáncer.

Las vacunas para el cáncer son otra forma de inmunoterapia actualmente en estudio. Las vacunas para las enfermedades infecciosas, tales como el sarampión, paperas, y el tétanos, son eficaces ya que exponen a las células inmunitarias del organismo a formas debilitadas de antígenos que están presentes en la superficie del agente infeccioso. Esta exposición hace que las células inmunitarias produzcan más células plasmáticas, que producen anticuerpos. Las células T que reconocen al agente infeccioso también se multiplican. Estas células T activadas recuerdan posteriormente la exposición. A la siguiente vez que el agente entra en el organismo, las células del sistema inmune ya están preparadas para responder y detener la infección.

Para el tratamiento del cáncer, los investigadores están desarrollando vacunas que pueden predisponer al sistema inmunitario del paciente para reconocer células cancerosas. Estas vacunas pueden ayudar al organismo a rechazar tumores y a evitar las recaídas del cáncer. A diferencia de las vacunas contra enfermedades infecciosas, las vacunas para el cáncer se diseñan para su inyección después de que diagnostique la enfermedad, en lugar de antes de que se desarrolle. Las vacunas para el cáncer administradas cuando el tumor es pequeño pueden ser capaces de erradicar el cáncer. Los ensayos clínicos de vacunas para el cáncer temprano (trabajos de investigación con personas) implican principalmente a pacientes con melanoma. Actualmente, también se están estudiando vacunas para el cáncer en el tratamiento de muchos otros tipos de cáncer, incluyendo linfomas y cánceres de riñón, de mama, de ovario, de próstata, de colon, y de recto. Los investigadores también están investigando formas de uso de las vacunas para el cáncer en combinación con otros MRB.

Al igual que otras formas de tratamiento para el cáncer, las terapias biológicas pueden provocar una serie de efectos secundarios, que pueden variar ampliamente de un paciente a otro. Pueden desarrollarse erupciones o inflamación en el sitio donde se inyectan los MRB. Varios MRB, incluyendo interferones e interleucinas, pueden provocar síntomas similares a la gripe, incluyendo fiebre, escalofríos, náuseas, vómitos, y pérdida de apetito. La fatiga es otro efecto secundario común de los MRB. También puede verse afectada la presión sanguínea. Los efectos secundarios de la IL-2 pueden ser a menudo graves, dependiendo de la dosis administrada. Se necesita controlar estrechamente al paciente durante el tratamiento. Los efectos secundarios de los CSF pueden incluir dolor de huesos, fatiga, fiebre, y pérdida de apetito. Los efectos secundarios de los MOAB varían, y pueden aparecer reacciones alérgicas graves. Las vacunas para el cáncer pueden provocar dolores musculares y fiebre.

Composiciones y métodos sensibilizantes - Dosificación, modo de administración y formulaciones farmacéuticas

La presente invención proporciona una composición que comprende un polinucleótido de la familia de SPARC que codifica el polipéptido de la familia de SPARC con una secuencia de SEC ID N°: 1 y un agente terapéutico. El polinucleótido de la familia de SPARC se proporciona en forma de una cantidad terapéuticamente eficaz para sensibilizar a una célula cancerosa o a un paciente al tratamiento por dicho agente terapéutico.

El agente terapéutico puede ser cualquier agente adecuado para una terapia adecuada, tal como se describe en el presente documento y tal como se conoce en la técnica. Puede ser un agente quimioterapéutico, es decir, un fármaco, por ejemplo, 5-fluorouracilo; puede ser un agente de radiación, tal como un anticuerpo radiomarcado, un radiosensibilizante, o un implante de semilla radiactiva. El agente terapéutico también puede ser un agente fotosensibilizante, tal como porfímero sódico; o un agente de terapia génica (por ejemplo, un vector), o puede ser un agente inmunoterapéutico, tal como una célula inmunitaria, un anticuerpo, o una citocina.

En el presente documento se divulga una composición que comprende un polipéptido de la familia de SPARC y una célula resistente a la quimioterapia.

En el presente documento se divulga una célula recombinante que comprende una región de control de la transcripción heteróloga asociada operativamente con un polinucleótido de la familia de SPARC.

Además de sensibilizar a una muestra o a un mamífero a la terapia para el cáncer, el uso de las composiciones objeto de la presente invención puede reducir la dosificación de una terapia, reduciendo de este modo los efectos secundarios causados por la terapia para el cáncer.

Las composiciones anteriores pueden ser una composición farmacéutica que incluye un vehículo o excipiente farmacéuticamente aceptable.

Tal como se usa en el presente documento, un "transportador" se refiere a cualquier sustancia adecuada como vehículo para administrar una CPA a un sitio de acción *in vitro* o *in vivo* adecuado. Como tal, los transportadores

pueden actuar como excipientes para la formulación de un reactivo terapéutico o experimental que contiene una CPA. Los transportadores preferidos son capaces de mantener una CPA en una forma que sea capaz de interactuar con una célula T. Los ejemplos de dichos transportadores incluyen, pero sin limitación, agua, solución salina tamponada con fosfato, suero salino, solución de Ringer, solución de dextrosa, soluciones que contienen suero, solución de Hank y otras soluciones fisiológicamente equilibradas acuosas o medio de cultivo celular. Los transportadores acuosos también pueden contener sustancias auxiliares adecuadas necesarias para aproximarse a las condiciones fisiológicas del receptor, por ejemplo, potenciación de la estabilidad química e isotonicidad. Las sustancias auxiliares adecuadas incluyen, por ejemplo, acetato de sodio, cloruro de sodio, lactato sódico, cloruro de potasio, cloruro de calcio, monolaurato de sorbitán, oleato de trietanolamina, y otras sustancias usadas para producir tampón fosfato, tampón Tris, y tampón bicarbonato.

Puede usarse una composición que comprende un polipéptido de la familia de SPARC y un agente terapéutico para sensibilizar a un cáncer *in vitro* poniendo en contacto directo a la muestra de cáncer con una cantidad eficaz del polipéptido de la familia de SPARC purificado. Puede administrarse a un mamífero (por ejemplo, un paciente con cáncer) una composición que comprende un polipéptido de la familia de SPARC para lograr el efecto de sensibilización *in vivo*. Además, una muestra de cáncer, ya sea células o tejido, puede obtenerse del mamífero y sensibilizarse usando el polipéptido de la familia de SPARC *ex vivo* antes de devolverse al mamífero.

Puede introducirse un polinucleótido de la familia de SPARC de la presente invención en una célula cancerosa *in vitro* para sensibilizar la respuesta de la célula cancerosa, o puede administrarse a un mamífero *in vivo* mediante un vector adecuado tal como se conoce en la técnica y tal como se describe anteriormente en el presente documento. Además, puede introducirse el vector *ex vivo* en células cancerosas o tejido obtenidos de un mamífero que lo necesite, y después devolverse las células o el tejido al mamífero que lo necesite. Al ser una proteína secretada, un polipéptido de la familia de SPARC producido por dichas células introducidas *ex vivo* puede funcionar en el ambiente local para sensibilizar no solo a las células modificadas, sino también a las células cancerosas vecinas no modificadas.

Puede introducirse una composición que comprende una célula recombinante en un mamífero para sensibilizarlo al tratamiento.

El tamaño de la dosis del sujeto, el número de dosis, la frecuencia de administración de la dosis, y el modo de administración pueden determinarse y optimizarse usando métodos conocidos en la técnica (véase, por ejemplo, Hardman *et al.*, Ceds 1995, Goodman & Gilman's, The Pharmacological Basis of Therapeutics, novena edición, McGraw-Hill).

Las dosificaciones de cada terapia para el tratamiento de diversos pacientes de cáncer se conocen en la técnica y pueden determinarse por un médico experto. Por ejemplo, una dosis de polipéptido de SPARC adecuada puede estar en el intervalo de 0,01 a 100 mg de polipéptido de SPARC por kilogramo de peso corporal del receptor por día, preferentemente en el intervalo de 0,2 a 10 mg por kilogramo de peso corporal por día. Puede administrarse un polipéptido de SPARC de la presente invención en una dosis adecuada en el intervalo de 0,01 a 100 mg de polinucleótido por kilogramo de peso corporal del receptor por día, preferentemente en el intervalo de 0,2 a 10 mg por kilogramo de peso corporal por día. Las células que comprenden un polinucleótido de SPARC recombinante pueden administrarse a una dosificación en el intervalo de 104-1010 por kilogramo de peso corporal del receptor, preferentemente en el intervalo de 106-108 por kilogramo de peso corporal del receptor. La dosis deseada se presenta preferentemente una vez al día, pero puede dosificarse como dos, tres, cuatro, cinco, seis o más subdosis administradas a intervalos adecuados a lo largo del día. Estas subdosis pueden administrarse en formas de dosis unitaria, por ejemplo, que contiene de 10 a 1500 mg, preferentemente de 20 a 1000 mg, y lo más preferentemente de 50 a 700 mg del polipéptido de la familia de SPARC por forma de dosificación unitaria. Las dosificaciones del polipéptido de la familia de SPARC o del polinucleótido de la familia de SPARC, o las células que comprenden un polinucleótido de la familia de SPARC útiles de acuerdo con la invención variarán dependiendo de la afección que vaya a tratarse o prevenirse y de la identidad del polipéptido o polinucleótido de la familia de SPARC que se esté usando. Puede efectuarse estimaciones de las dosis eficaces y de las semividas *in vivo* para la composición individual abarcada por la invención basándose en pruebas *in vivo* usando un modelo animal, tal como el modelo de ratón descrito en el presente documento o una adaptación de dicho método para animales mayores.

Aplicaciones *in vitro/ex vivo*

Las composiciones proporcionadas por la presente invención pueden usarse para sensibilizar a una célula cancerosa *in vitro* usando métodos conocidos en la técnica, y tal como se describen anteriormente en el presente documento. Por lo tanto, la presente invención proporciona un método para sensibilizar a una célula cancerosa a un tratamiento terapéutico, comprendiendo el método poner en contacto la muestra de sangre con una cantidad eficaz de una composición que comprende un polipéptido de la familia de SPARC o de un polinucleótido que codifica un polipéptido de la familia de SPARC. En el presente documento se divulga un método para sensibilizar *ex vivo* a un mamífero al que se le ha diagnosticado cáncer a un tratamiento terapéutico, comprendiendo el método: (1) obtener una muestra de cáncer del mamífero; (2) poner en contacto la muestra de cáncer con una cantidad eficaz de una

composición que comprende un polipéptido de la familia de SPARC o un polinucleótido que codifica un polipéptido de la familia de SPARC; y (3) devolver la muestra de cáncer después de la puesta en contacto de (2) al mamífero.

La terapia génica *ex vivo* se refiere al aislamiento de células de un animal, a la administración de un polinucleótido a las células, *In vitro*, y a la devolución de las células modificadas de nuevo al animal. Esto puede implicar la extracción quirúrgica de tejido/órganos de un animal o el cultivo primario de células y tejidos. Anderson *et al.*, Patente de Estados Unidos n.º 5.399.346 divulgan métodos terapéuticos *ex vivo*. Este método es aplicable debido a que el polipéptido de la familia de SPARC es un polipéptido secretado. La devolución de las células modificadas de nuevo al mamífero puede aumentar la concentración extracelular de un polipéptido de la familia de SPARC de manera local y de este modo sensibilizar a las células cancerosas no modificadas que están en proximidad a las células modificadas.

Cuando se necesita extraer muestras de tejido de un mamífero para aplicación *ex vivo*, pueden prepararse extractos celulares a partir de biopsias de tejido de pacientes incluyendo, pero sin limitación, de cerebro, corazón, pulmón, ganglios linfáticos, ojos, articulaciones, piel y neoplasias asociadas a estos órganos. La "biopsia de tejido" también abarca la recogida de fluidos biológicos, incluyendo, pero sin limitación, sangre, plasma, esputo, orina, fluido cefalorraquídeo, lavados, y muestras de leucoforesis. Tal como se enseña en el presente documento, las "biopsias de tejido" pueden extraerse de tumores de mama, ovario o próstata. Las "biopsias de tejido" se obtienen usando técnicas bien conocidas en la técnica, incluyendo aspiración con aguja y biopsia por punción de la piel.

En general, cuando se introduce un polinucleótido en células en cultivo (por ejemplo, mediante una de las técnicas de transfección descritas anteriormente) solo una pequeña porción de las células (aproximadamente 1 de cada 10^5) integran típicamente el polinucleótido transfectado en sus genomas (es decir, el polinucleótido se mantiene en la célula de manera episómica). Por lo tanto, para identificar células que han captado polinucleótido exógeno, es ventajoso transfectar polinucleótidos que codifiquen un marcador de selección en la célula junto con el polinucleótido (o los polinucleótidos) de interés, es decir, un polinucleótido de la familia de SPARC, tal como se ha descrito anteriormente en el presente documento.

Aplicaciones *in vivo*

La composición proporcionada puede administrarse a un mamífero, por ejemplo, en un método para sensibilizar para un tratamiento terapéutico. En el presente documento se divulga un método para sensibilizar para un tratamiento terapéutico a un animal *in vivo* al que se le ha al que se le ha diagnosticado cáncer, comprendiendo el método administrar al mamífero una cantidad eficaz de una composición que comprende un polipéptido de la familia de SPARC o un polinucleótido que codifica un polipéptido de la familia de SPARC.

El modo de administración de una composición de la presente invención puede depender del propósito concreto para la administración, de la salud general y de la afección del paciente y del juicio del médico o técnico que administre el vehículo diana. Puede administrarse una composición de la presente invención a un animal usando una diversidad de métodos. Dichos métodos de administración pueden incluir administración parenteral, tópica, oral o local, tal como por vía intradérmica. Puede administrarse una composición en una diversidad de formas de dosificación unitaria, dependiendo del método de administración. Los métodos de administración preferidos para una composición de la presente invención incluyen administración intravenosa, administración local (por ejemplo, intratumoral) mediante, por ejemplo, inyección, inyección intradérmica, inyección intramuscular, inyección intraperitoneal e inhalación. Para modos de administración particulares, las composiciones divulgadas en el presente documento pueden formularse en un excipiente. Puede administrarse una composición a mamíferos, y más preferentemente a seres humanos.

Inyección: Las formas farmacéuticas adecuadas para uso inyectables inducen soluciones acuosas o dispersiones estériles para la preparación extemporánea de soluciones o dispersiones inyectables estériles. En todos los casos, la forma tiene que ser estéril y tiene que ser fluido hasta el punto de que sea fácilmente inyectable. Tiene que ser estable en las condiciones de fabricación y almacenamiento y tienen que conservarse contra la acción contaminante de los microorganismos, tales como bacterias y hongos. El transportador puede ser un disolvente o medio de dispersión que contiene, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol, y polietilenglicol líquido, y similares), mezclas adecuadas de los mismos, y aceites vegetales. Puede mantenerse la fluidez adecuada, por ejemplo, mediante el uso de un recubrimiento, tales como lecitina, manteniendo el tamaño de partícula necesario en el caso de las dispersiones y mediante el uso de tensioactivos. La prevención de la acción de microorganismos puede lograrse por medio de varios agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido sórbico, timerosal, y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares o cloruro sódico. Puede lograrse la absorción prolongada de las composiciones inyectables mediante el uso en las composiciones de agentes retardantes de la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Las soluciones inyectables estériles se preparan incorporando los principios activos en la cantidad necesaria en el disolvente adecuado con varios de los otros ingredientes indicados anteriormente, según sea necesario, seguido de esterilización por filtración. En general, las dispersiones se preparan incorporando los varios principios activos

esterilizados en un vehículo estéril que contiene el medio de dispersión básico y los demás ingredientes necesarios de entre los enumerados anteriormente. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los métodos preferidos de preparación son las técnicas secado al vacío y criodesecación que producen un polvo del principio activo más cualquier otro ingrediente adicional deseado a partir de una solución previamente esterilizada por filtración del mismo.

Administración oral: Las composiciones farmacéuticas para administración oral se formulan usando vehículos farmacéuticamente aceptables bien conocidos en la técnica en dosificaciones adecuadas para administración oral. Dichos vehículos permiten que las composiciones farmacéuticas se formulen como comprimidos, píldoras, grageas, cápsulas, líquidos, geles, jarabes lechados, suspensiones y similares, para su ingestión por parte del paciente.

Las preparaciones farmacéuticas para uso oral se obtienen mediante una combinación de compuestos activos con excipiente sólido, triturando opcionalmente una mezcla resultante, y procesando la mezcla de gránulos, después de añadir agentes auxiliares adecuados, si se desea, para obtener comprimidos o núcleos de grageas. Los excipientes adecuados son cargas de carbohidratos o proteínas, tales como azúcares, incluyendo lactosa, sacarosa, manitol, o sorbitol; almidón de maíz, trigo, arroz, patata, u otras plantas; celulosa, tal como metilcelulosa, hidroxipropilmetilcelulosa, o carboximetilcelulosa de sodio; y gomas, tales como goma arábiga y tragacanto; y proteínas, tales como gelatina y colágeno. Si se desea, pueden añadirse agentes disgregantes o solubilizantes, tales como polivinilpirrolidona reticulada, agar, ácido algínico, o una sal del mismo, tal como alginato de sodio.

Los núcleos de grageas se proporcionan con recubrimientos adecuados, tales como soluciones de azúcar concentradas, que también pueden contener goma arábiga, talco, polivinilpirrolidona, gel de carbopol, polietilenglicol, y/o dióxido de titanio, soluciones de laca, y disolventes orgánicos adecuados o mezclas de disolventes. Pueden añadirse colorantes o pigmentos a los comprimidos o recubrimientos de grageas para la identificación del producto o para caracterizar la cantidad del principio activo, es decir, la dosis.

Las preparaciones farmacéuticas que se usan por vía oral incluyen cápsulas de ajuste a presión hechas de gelatina, así como cápsulas blandas selladas hechas de gelatina y un recubrimiento, tal como glicerol o sorbitol. Las cápsulas de ajuste a presión pueden contener ingredientes activos mezclados con una carga o aglutinantes, tales como lactosa o almidones, lubricantes, tales como talco o estearato de magnesio y, opcionalmente, estabilizantes. En las cápsulas blandas, los compuestos activos pueden disolverse o suspenderse en líquidos adecuados, tales como aceites grasos, parafina líquida, o polietilenglicol líquido con o sin estabilizantes.

Administración nasal: Para administración nasal, se usan en la formulación penetrantes adecuados para la barrera concreta a permear. Dichos penetrantes se conocen generalmente en la técnica.

Uso subcutáneo e intravenoso: Para uso subcutáneo e intravenoso, la composición de la invención se proporcionará generalmente en soluciones o suspensiones acuosas estériles, tamponadas a un pH e isotonicidad adecuados. Los vehículos acuosos adecuados incluyen solución de Ringer y cloruro de sodio isotónico. Las suspensiones acuosas de acuerdo con la invención pueden incluir agentes de suspensión, tales como derivados de celulosa, alginato de sodio, polivinilpirrolidona y goma de tragacanto, y un agente humectante, tal como lecitina. Los conservantes adecuados para suspensiones acuosas incluyen p-hidroxibenzoato de etilo y de n-propilo.

La composición útil, tal como se describe en el presente documento, también puede presentarse como formulaciones de liposomas.

La terapia génica usando las composiciones proporcionadas por la presente invención puede llevarse a cabo de acuerdo con métodos generalmente aceptados, por ejemplo, tal como se describe por Friedman en "Therapy for Genetic Disease," T. Friedman, ed., Oxford University Press (1991), págs. 105-121.

Las composiciones farmacéuticas de la presente invención pueden fabricarse de un modo conocido en la técnica, por ejemplo, mediante procesos convencionales de mezclado, disolución, granulación, formación de grageas, levigación, emulsión, encapsulación, atrapamiento o liofilización.

Después de que se hayan preparado las composiciones farmacéuticas que comprenden un agente terapéutico de la invención en un transportador aceptable, se colocan en un contenedor adecuado y se etiquetan para el tratamiento de una afección indicada con información que incluye la cantidad, frecuencia y método de administración.

La dosis exacta se selecciona por el médico individual a la vista del paciente que se va a tratar. La dosificación y la administración se ajustan para proporcionar niveles suficientes del resto activo o para mantener el efecto deseado. Los factores adicionales que pueden tenerse en cuenta incluyen la gravedad de la patología (por ejemplo, localización de la enfermedad, la edad, peso, y género del paciente, de la dieta, tiempo y frecuencia de administración, combinación (o combinaciones) de fármacos, reacciones de sensibilidad, y tolerancia/respuesta a la terapia). Pueden administrarse composiciones farmacéuticas de acción prolongada cada 3 a 4 días, cada semana, o una vez cada dos semanas dependiendo de la semivida y de la velocidad de eliminación de la formulación particular. Se proporciona orientación general en cuanto a las dosis particulares y a los métodos de administración para otras

aplicaciones en la bibliografía (véanse las Patentes de Estados Unidos n.º 4.657.760; 5.206.344; y 5.225.212, incorporadas al presente documento por referencia). Los expertos en la materia emplearán típicamente diferentes formulaciones para oligonucleótidos y vectores de terapia génica que para proteínas o sus inhibidores. De manera similar, la administración de los polinucleótidos o polipéptidos será específica para células, afecciones y localizaciones concretas y similares.

La composición puede formularse tal como se describe en la Patente de Estados Unidos n.º 6187330 que proporciona una composición para la liberación controlada de un péptido o proteína que comprende un polímero biocompatible, bioerosionable que tiene dispersado en el mismo una fase de matriz vítrea que comprende el péptido o proteína y un termoprotector, teniendo dicha fase de matriz vítrea una temperatura de transición vítrea por encima del punto de fusión del polímero. Ya que el fármaco de péptido o proteína es estable dentro de la composición, puede formarse convenientemente en su estado fundido, en dispositivos de forma adecuada para su uso como implantes de administración de fármaco, por ejemplo, en forma de varillas, películas, perlas u otras formas deseadas.

Determinación de la resistencia o sensibilidad a un tratamiento terapéutico

La determinación de una muestra de cáncer (por ejemplo, células o tejido) que responde a un tratamiento terapéutico puede llevarse a cabo mediante cualquier método conocido en la técnica. Por ejemplo, mediante una prueba de resistencia a fármacos en cultivo celular (CCDRT). CCDRT se refiere a ensayar una muestra de cáncer (por ejemplo, extraída de un paciente mamífero) en el laboratorio para fármacos que pueden usarse para tratar el cáncer del paciente. La prueba puede identificar a qué fármacos es sensible la muestra y a cuáles es resistente, lo que indica qué fármacos tienen mayores probabilidades de funcionar y qué fármacos tienen menos probabilidades de funcionar en el paciente. La sensibilidad de la muestra de cáncer (por lo tanto, del paciente) puede detectarse mediante el uso del polinucleótido de SPARC tal como se ha divulgado en el presente documento. La CCDRT puede llevarse a cabo de nuevo con el tratamiento sensibilizante y determinar si una composición sensibilizante proporcionada por la presente invención puede sensibilizar la respuesta de la muestra de cáncer a un tratamiento específico, o no. Puede decirse que una composición sensibiliza a un tratamiento terapéutico si la respuesta medida mediante CCDRT aumenta en al menos un 20%, por ejemplo, 30%, 40%, 50%, 80%, 100% (2 vez), o 3 veces, 4 veces, 5 veces, o más cuando se compara con la respuesta en ausencia de la composición sensibilizante.

La CCDRT incluye típicamente ensayos de proliferación celular y ensayos de muerte celular.

El ensayo de proliferación celular mide la proliferación de células. Puede efectuarse con el ensayo de incorporación de timidina radiactiva descrito originariamente por Tanigawa y Kern (anteriormente citado). En este ensayo, aplicado solamente a tumores sólidos y no a neoplasias hematológicas, se cultivan las células tumorales suspendidas en agarosa blanda durante 4-6 días en presencia continua de fármacos antineoplásicos. Al final del periodo de cultivo, se introduce timidina radioactiva y se comparan las diferencias en la incorporación de timidina homóloga en el ADN entre los cultivos de control y tratados con fármaco. Kern y Weisenthal analizaron los datos de correlación clínica y definieron el concepto de "resistencia a fármacos extrema" o EDR [Kern DH, Weisenthal LM. J Natl Cancer Inst 1990; 82: 582-588]. Esto se definió como un resultado de ensayo que era una desviación estándar más resistente que la mediana del resultado para ensayos de comparación de bases de datos. Los pacientes tratados con agentes individuales que mostraron EDR en el ensayo virtual nunca disfrutaron de una respuesta parcial o completa. Kern y Weisenthal también definieron la "resistencia a fármacos baja" (LDR) como un resultado menos resistente que la mediana y la "resistencia a fármacos intermedia" (IDR) como un resultado más resistente que la mediana, pero menos resistente que EDR (en otras palabras, entre la mediana y una desviación estándar más resistente que la mediana).

Los principios y los datos de correlación clínica con el ensayo "EDR" de timidina se revisaron en 1992 (Weisenthal LM, Kern DH. Oncology (U S A) 1992; 5: 93-103]. Desde aquel momento solo se han publicado unos pocos estudios de seguimiento. Uno de dichos estudios demostró que la EDR a uno o más de los agentes individuales usados en una combinación de dos fármacos no se asocia con una probabilidad menor de respuesta a la combinación de los dos fármacos en el escenario de quimioterapia intraperitoneal de los cánceres de apéndice y de colon (Femández-Trigo V, Shamsa F, Vidal-Jove J, Kern DH, Sugarbaker PH. Am J Clin Oncol 1995; 18: 454-460). Es, sin embargo, posible que la respuesta a las altas concentraciones de fármacos obtenibles con la quimioterapia intraperitoneal pueda asociarse más estrechamente con la penetración del fármaco en el tumor que a la resistencia a fármacos intrínseca de las células tumorales. También se demostró que la EDR al paclitaxel no parece ser un factor pronóstico en pacientes con cáncer de ovario o en pacientes con carcinoma peritoneal primario tratados con paclitaxel más platino (Eltabbakh GH, Piver MS, Hempling RE, *et al.* Gynecol Oncol 1998; 70: 392-397; Eltabbakh GH. J Surg Oncol 2000; 73: 148-152). Sin embargo, se ha comunicado recientemente que la EDR al platino en el cáncer de ovario puede tener implicaciones de pronóstico (Fruenhaus, J., *et al.* Proc ASCO, v.20, Abs 2529, 2001). También se comunicó que los pacientes de cáncer de mama no tratados previamente con tumores que mostraban LDR (definida anteriormente) tienen tiempos superiores hasta la progresión y supervivencias generales que los pacientes que muestran bien IDR o EDR (Mehta, R.S., *et al.* Breast Cancer Res Treat 66:225-37, 2001).

El ensayo de "EDR" de timidina tiene una elevada especificidad (>98%) para la identificación de agentes individuales inactivos, pero una baja sensibilidad (<40%). En otras palabras, se predice que un fármaco con "EDR" definida en ensayo es prácticamente seguro que sea inactivo como agente individual (alta especificidad para identificar fármacos inactivos), pero muchos fármacos sin "EDR" también serán inactivos (baja sensibilidad para identificar fármacos inactivos).

Una segunda forma de ensayo de proliferación celular es el sistema de cultivo de células tumorales adherentes, basado en la comparación del crecimiento de monocapas de células sobre una "matriz adherente celular" de diseño patentado (Ajani JA, Baker FL, Spitzer G, *et al.* J Clin Oncol 1987). También se describieron correlaciones clínicas positivas en esta publicación.

Tal como se enseña en el presente documento, pueden usarse ensayos de formación de colonias para medir la proliferación celular. En esta prueba, se cultivan células *in vitro* en agar blando (medio de cultivo tisular que contiene agar como agente gelificante; también citado como agar semi-sólido) u otro medio altamente viscoso, que contiene, por ejemplo, metilcelulosa, gel de plasma o coágulos de fibrina. Estos medios semisólidos reducen el movimiento celular y permiten que las células individuales se desarrollen en clones celulares que se identifican como colonias individuales. Estos ensayos también se citan generalmente como ensayos donogénicos. Los ensayos de formación de colonias se conocen bien en la técnica, por ejemplo, véase Rizzino, A Soft agar growth assays for transforming growth factors and mitogenic peptides. Methods in Enzymology 146: 341-53 (1987) y en algunas realizaciones, la apoptosis se mide mediante un ensayo de marcado del extremo libre por dUTP mediado por desoxinucleotidil transferasa terminal (TUNEL) que se conoce bien en la técnica y del que hay materiales y métodos disponibles como material en línea de soporte en Science Online.

En oposición a medir la proliferación celular, hay una familia de ensayos estrechamente relacionados basados en el concepto de eliminación celular total, o en otras palabras, la muerte celular que sucede en la población de células tumorales (Weisenthal LM, Shoemaker RH, Marsden JA, Dill PL, Baker JA, Moran EM. Recent Results Cancer Res 1984; 94: 161-173; Weisenthal LM, Lippman ME. Cancer Treat Rep 1985; 69: 615-632; Weisenthal LM. Cell culture assays for hematologic neoplasms based on the concept of total tumor cell kill. En: Kaspers GJL, Pieters R, Twentymann PR, Weisenthal LM, Veerman AJP, eds. Drug Resistance in Leukemia and Lymphoma. Langhorne, PA: Harwood Academic Publishers, 1993: 415-432; Weisenthal LM. Contrib Gynecol Obstet 1994; 19: 82-90). Los conceptos subyacentes a los ensayos de muerte celular son relativamente simples, aunque las características técnicas y la interpretación de los datos pueden ser muy complejas.

Los conceptos tecnológicos básicos son sencillos. Por ejemplo, se obtiene un espécimen fresco de una neoplasia viable. Con más frecuencia, el espécimen es un espécimen quirúrgico de un tumor sólido viable. Con menor frecuencia, es una efusión maligna, médula ósea, o un espécimen de sangre periférica que contiene células "tumorales" (una palabra usada para describir células de una neoplasia sólida o hematológica). Estas células se aíslan y después se cultivan en la presencia continua o ausencia de fármacos, con mayor frecuencia de 3 a 7 días. Al final del periodo de cultivo, se efectúa una medición de lesiones celulares, que se correlaciona directamente con la muerte celular. Hay pruebas de que la mayoría de los fármacos anticáncer disponibles pueden funcionar mediante un mecanismo que causa un daño suficiente como para desencadenar la denominada muerte celular programada, o apoptosis (Hickman JA. Cancer Metastasis Rev 1992; 11: 121-139; Zunino F, Perego P, Pilotti S, Pratesi G, Supino R, Arcamone F. Pharmacol Ther 1997; 76: 177-185).

Aunque hay métodos para medir específicamente la apoptosis que pueden aplicarse a células cultivadas, hay dificultades prácticas para aplicar estos métodos a poblaciones mixtas (y agrupadas) de células tumorales y células normales. Por lo tanto, se han aplicado mediciones más generales de muerte celular. Estas incluyen: (1) pérdida retrasada de la integridad de la membrana (que se ha hallado como un subrogado útil para la apoptosis), medida mediante tinción diferencial en el método de ensayo *DISC*, que permite reconocer los efectos selectivos de fármacos contra células tumorales en una población mixta de células tumorales y normales (Weisenthal LM, Kern DH. Oncology (EE.UU.) 1992; 5: 93-103; Weisenthal LM, Marsden JA, Dill PL, Macaluso CK. Cancer Res 1983; 43: 749-757), (2) pérdida de la actividad del ciclo de Krebs mitocondrial, medida mediante el ensayo *MTT* (Carmichael J, DeGraff WG, Gazdar AF, Minna JD, Mitchell JB. Cancer Res 1987; 47: 936-942), (3) pérdida de ATP celular, medida en el ensayo de *ATP* (Kangas L, Gronroos M, Nieminen AL. Med Biol 1984; 62: 338-343; Garewal HS, Ahmann FR, Schiffman RB, Celniker A. J Natl Cancer Inst 1986; 77: 1039-1045; Sevin B-U, Peng ZL, Perras JP, Ganjei P, Penalver M, Averette HE. Gynecol Oncol 1988; 31: 191-204), y (4) pérdida de la actividad de esterasa de la membrana celular y de la integridad de la membrana celular, medida mediante el ensayo de *diacetato de fluoresceína* (Rotman B, Teplitz C, Dickinson K, Cozzolino JP. In vitro Cell Dev Biol 1988; 24: 1137-1138; Larsson R, Nygren P, Ekberg M, Slater L. Leukemia 1990; 4: 567-571; Nygren P, Kristensen J, Jonsson B, *et al.* Leukemia 1992; 6: 1121-1128).

Tal como se divulga en el presente documento, la apoptosis puede medirse mediante un ensayo de marcado del extremo libre por dUTP mediado por desoxinucleotidil transferasa terminal (TUNEL) que se conoce bien en la técnica y por ejemplo, tal como se describe en materiales y métodos disponibles como material en línea de soporte en Science Online.

Además, la sensibilidad o la resistencia de un animal a un tratamiento puede determinarse directamente midiendo el tamaño del tumor antes y después del tratamiento y/o durante un periodo de tiempo del tratamiento. Si el tamaño del tumor disminuye en un 50%, preferentemente un 75%, más preferentemente un 85%, lo más preferentemente un 100% con un tratamiento, entonces se dice que el animal es sensible (no resistente) al tratamiento. Por otra parte, se considera que el animal es resistente al tratamiento si el tamaño del tumor se reduce en al menos un 25%, preferentemente un 50%, más preferentemente un 75%, lo más preferentemente un 100% después de administrar una composición sensibilizante al tratamiento de la presente invención en comparación con el tumor después del tratamiento pero en ausencia de la una composición de la presente invención, diciéndose entonces que la composición es eficaz para sensibilizar al tratamiento en el animal. En seres humanos, el tamaño del tumor puede compararse durante un periodo ventana de 6 meses de tratamiento, en otros animales, este periodo ventana varía, por ejemplo puede usarse una ventana de 4-6 semanas para ratones. Se entiende que el periodo ventana real para comparar el tamaño del tumor puede determinarse de acuerdo con el conocimiento de la técnica y del tumor particular que se va a tratar.

Además, también puede determinarse si una célula es resistente a un tratamiento midiendo la expresión de un polipéptido o polipéptido de la familia de SPARC tal como se ha descrito anteriormente en el presente documento. Por lo tanto, la presente invención proporciona un método para evaluar una primera muestra de cáncer respecto de su resistencia a un tratamiento terapéutico, que comprende: (a) medir el nivel de expresión de un ARNm o polipéptido de la familia de SPARC, o el nivel extracelular de un polipéptido de la familia de SPARC en la primera muestra de cáncer; (b) medir el nivel de expresión del ARNm o polipéptido de la familia de SPARC, o el nivel extracelular del polipéptido de la familia de SPARC en una segunda muestra de cáncer que no muestra resistencia al tratamiento terapéutico; (c) comparar los niveles de expresión o los niveles extracelulares obtenidos en (a) y (b), en el que un nivel menor de expresión o nivel extracelular en (a) es indicativo de que la primera muestra de cáncer es resistente al tratamiento terapéutico.

Además, en el presente documento se enseña un método para identificar un agente que modula la expresión de un ARNm o polipéptido de la familia de SPARC, o la secreción de un polipéptido de la familia de SPARC, que comprende: (a) medir el nivel de expresión del ARNm o polipéptido de la familia de SPARC, o el nivel extracelular del polipéptido de la familia de SPARC en una muestra; (b) poner en contacto un agente candidato con la muestra; (c) después de la puesta en contacto de (b), medir la expresión o el nivel extracelular del ARNm o polipéptido de la familia de SPARC, o el nivel extracelular del polipéptido de la familia de SPARC en la muestra de (b); (d) comparar los niveles de expresión o los niveles extracelulares de (a) y (c), en el que un nivel de expresión o un nivel de expresión diferencial en (a) y (c) indica que el agente candidato es un agente que modula la expresión del ARNm o del polipéptido de la familia de SPARC, o de la secreción del polipéptido de la familia de SPARC.

Los niveles de expresión de un polipéptido o de un polinucleótido de la familia de SPARC y los niveles de secreción de un polipéptido de SPARC pueden medirse tal como se describe anteriormente en el presente documento y mediante cualquier método conocido en la técnica.

Un agente, que potencia la expresión o secreción de un miembro de la familia de SPARC, puede usarse solo como un agente sensibilizante a la terapia, tal como se describe en la presente invención. El agente puede ser un químico, o una molécula biológica (por ejemplo, una proteína, o un polinucleótido, etc.).

Modelos animales

Los efectos terapéuticos de las composiciones proporcionadas pueden ensayarse en diversos modelos animales. Esto puede efectuarse *in vitro*, *ex vivo*, o *in vivo* tal como se ha descrito anteriormente en el presente documento.

Los modelos de ratón para trastornos proliferativos se conocen en la técnica y pueden encontrarse, por ejemplo, en la base de datos de ratones de Jackson Laboratory en la página de Internet www.jax.org y en el catálogo de The Jackson Laboratory, Jax-Mice- Junio de 2001- Mayo de 2003, o Jackson-Grusby L. 2002, Oncogene. 12;21(35):5504-14; Ghebranious N, Donehower LA, 1998, Oncogene. 24;17(25):3385-400; Palapattu GS, Bao S, Kumar TR, Matzuk MM. 1998, Cancer Detect Prev. 22(1):75-86). Por ejemplo, los modelos de ratón de tumor incluyen aquellos usados para el estudio de la leucemia mielógena crónica (LMC), defectos en moléculas de adhesión, genes que regulan el crecimiento y la proliferación, factores de crecimiento/receptores/citocinas, incidencia aumentada de tumores, oncogenes, toxicología y genes supresores de tumores.

Ejemplos

La invención se basa en la observación de que se encontró que SPARC estaba significativamente subexpresado en células resistentes a la quimioterapia, que el polipéptido de SPARC sensibiliza a células resistentes a la terapia para el cáncer al tratamiento del cáncer, que el ADN que codifica a SPARC sensibiliza a las células a la terapia para el cáncer, y que los animales con injertos de células transfectantes para SPARC muestran una reducción dramática en el crecimiento tumoral en comparación con animales a los que se injerta un control.

Ejemplo 1. MATERIALES Y MÉTODOS

Cultivo celular -Se mantuvo a la línea celular MIP-101 colorrectal en medio Eagle modificado de Dulbecco (DMEM) (Invitrogen) suplementado con suero bovino fetal (FBS) al 10% (Invitrogen), penicilina/estreptomina al 1% (Invitrogen) a 37°C y CO₂ al 5%. Se desarrollaron células MIP101 resistentes después de una incubación de larga duración con concentraciones en aumento de 5-fluorouracilo (5-FU), irinotecán (CPT-11), cisplatino (CIS), y etopósido (ETO). Las células MIP101 estables transducidas con SPARC (MIP/SP) se mantuvieron en DMEM suplementado con FBS al 10%, penicilina/estreptomina al 1%, y zeocina al 0,1% a 37°C y CO₂ al 5%.

Retrotranscripción analítica - Reacción en cadena de la polimerasa - Se extrajo el ARN total de células cultivadas (2 x 10⁶ células, confluencia al 75%) usando reactivo TRIZOL (Invitrogen) de acuerdo con el protocolo del fabricante. La RT-PCR se llevó a cabo usando un kit comercialmente disponible (BD Biosciences) siguiendo el protocolo del fabricante usando 1 µg de ARN total. Los cebadores específicos usados para amplificar SPARC: 5'-CGA AGA GGA GGT GGT GGC GGA AA-3' (con sentido) (SEC ID N°: 78) y 5'-GGT TGT TGT CCT CAT CCC TCT CAT AC-3' (antisentido) (SEC ID N°: 79). GAPDH: 5'-CTC TCT GCT CCT CCT GTT CGA CAG-3' (con sentido) (SEC ID N°: 80) y 5'-AGG GGT CTT ACT CCT TGG AGG CCA-3' (antisentido) (SEC ID N°: 81) se usaron como control interno y para normalizar los niveles de expresión génica. Se usaron los siguientes ajustes para la reacción: 50°C x 1 h, seguido de 37 ciclos de 94°C x 1 min, 65°C x 1 min, 72°C x 2 min, seguido de 72°C x 10 min e incubación a 4°C. Los productos de la PCR se separaron en gel de agarosa al 1% en tampón TAE (teñido con $\mu\text{g/ml}$ de bromuro de etidio) mediante electroforesis durante 1 h a 100 V y posteriormente se fotografiaron.

Cuantificación de la apoptosis - Para el ensayo TUNEL, se sembraron las células en cubreobjetos de vidrio en placas de 6 pocillos a 250.000 células/placa durante toda la noche antes de cualquier estudio de inducción pasadas 24 horas. Para la evaluación de la apoptosis después de SPARC exógeno (Haematologic Technologies Inc.), las células se incubaron con 5 µg/ml de SPARC durante 24 horas seguido de una exposición de 12 h a 1000 µM de 5-FU. Después se procesaron las células para su marcado con el kit de detección de apoptosis (Promega) de acuerdo con las instrucciones del fabricante. Para la cuantificación de la apoptosis, se sembraron las células a 250.000 células/pocillo en placas de 6 pocillos durante toda la noche, seguido de una incubación de 12 h con los siguientes agentes quimioterapéuticos: 5-FU 1000 µM, CPT-11 200 µM, cisplatino 100 µM, y etopósido 10 µM. Las células se recogieron usando un medio de disociación celular no enzimático (Sigma), se lavaron con suero salino tamponado con fosfato y posteriormente se tiñeron con anexina V y yoduro de propidio usando un kit de detección de apoptosis (R&D Research) de acuerdo con el protocolo del fabricante. La proporción de células marcadas con anexina V y yoduro de propidio se analizó mediante el analizador de citometría de flujo XL. Se recogieron datos de 100.000 sucesos.

Transfección y selección de clones - Se clonó el ADNc de SPARC en el vector de expresión pcDNA3.1. Las transfecciones se efectuaron con 2,0 µg del gen/construcción de vector de expresión usando el método de polietilenimina de Boussiff *et al.* (1995) con modificaciones menores (Tai *et al.*, 2002). Después de la transfección, se lavaron las células con suero salino tamponado con fosfato (PBS, pH 7,4) y se mantuvieron en medio de cultivo durante 24 horas, seguido de un cambio a un medio de selección adecuado que contenía zeocina al 1%. Las células se seleccionaron basándose en su resistencia a la zeocina y se propagaron las colonias y los clones individuales de estas colonias para su verificación posterior. Los clones transducidos de manera estable (MIP/SP) se exploraron respecto de la expresión de ARNm de SPARC mediante análisis de reacción en cadena de la polimerasa transcriptasa inversa (RT-PCR). Los clones MIP/SP con la mayor expresión de ARNm (mediante RT-PCR) y proteína (mediante transferencia de Western) de SPARC se detectaron para estudios *in vitro* e *in vivo* posteriores. Las líneas celulares de control usadas para este estudio incluyeron células MIP101 transducidas de manera estable solo con el vector pcDNA3.1 vacío (MIP/Zeo) y se seleccionaron basándose en su resistencia a la zeocina.

Análisis de transferencia de Western - Se extrajo la proteína total de las líneas celulares cultivadas en placas de 10 cm usando tampón de lisis CHAPS. Se sometieron a electroforesis 10-30 µg de proteína total usando SDS-PAGE y se transfirieron a una membrana de PVDF. Después de bloquear con una solución de leche sin grasa al 5%, se incubaron las membranas con anticuerpo anti-SPARC (1:1000, Haematologic Technologies Inc.) durante toda la noche a 4°C. La membrana se incubó posteriormente con anticuerpo secundario de conejo anti-ratón conjugado a HRP (1:2000) durante 1 h a temperatura ambiente y se detectó mediante el kit de quimioluminiscencia Extend-Dura (Pierce). La misma membrana se decapó usando el tampón de decapado para transferencia de Western Restore (Pierce) y posteriormente se volvió a sondear para tubulina con un anticuerpo primario de ratón anti-tubulina (Sigma) y un anticuerpo secundario de conejo anti-ratón conjugado a HRP (1:2000) como control interno.

Inmunohistoquímica - Las secciones en parafina de cánceres colorrectales o epitelio colónico normal fueron proporcionadas amablemente por el Dr. Maximo Loda (Dana Farber Cancer Institute, Boston). Antes de la tinción, se lavaron las secciones con suero salino tamponado con Tris (TBS) al 0,1% que contenía Triton X-100 al 0,1% (Sigma), se trataron con H₂O₂ al 1% durante 30 min, se lavaron en TBS/Triton al 0,1% durante 30 min (x 3) a temperatura ambiente, y se bloquearon con BSA al 3% en TBS/Triton al 0,1% durante 1 h. Entonces se incubaron las secciones con anticuerpo de ratón anti-SPARC (1:50) durante toda la noche a 4°C (Haematologic Technologies Inc.), se lavaron varias veces con TBS/Triton y se contratiñeron con solución compleja de avidina-biotina-peroxidasa

(ABC) (kits de ABC Vecstain, Vector Laboratories Inc, Burlingame, CA) durante 1 h, seguido de incubación en solución DAB. Las secciones se montaron usando Pmount.

Ensayo de formación de colonias - Para los estudios de supervivencia clonogénicos, se sembraron células MIP101 progenitoras y células MIP/SP a 1000 células/placa en placas de 48 pocillos y se incubaron con concentraciones crecientes de 5-FU (0, 10 μ M, 100 μ M, 1000 μ M), CPT-11 (0, 1 μ M, 10 μ M, 100 μ M), o etopósido (0, 10 μ M, 100 μ M, 1000 μ M) durante 4 días. Entonces se lavaron las células con PBS y se incubaron en medio reciente que contenía las concentraciones adecuadas de quimioterapia durante 7 días adicionales. Se tiñó cada pocillo con violeta cristal y se contaron las colonias con más de 50 células. El número de colonias formadas en el grupo tratado se calculó basándose en las colonias formadas respecto de las células de control no tratadas.

Sobrenadante concentrado que contiene SPARC [SPARC] - Se sembraron células MIP/SP a 1×10^6 células en matraces de 100 cm en DMEM (FBS al 10%, penicilina/estreptomicina al 1%, zeocina al 0,1%) durante 24 horas. Las células se lavaron posteriormente con PBS dos veces y se incubaron en medio VP-SFM sin suero suplementado con glutamina 4 mM (Invitrogen) durante 72 horas. Este medio se concentró de 500 ml a 2 ml usando unidades de filtro Centricon (Millipore) a 4°C. Todos los medios recogidos y procesados mediante este método se usaron para los estudios posteriores en animales.

Estudios en animales - Se usaron modelos en animales de xenoinjerto de tumores para evaluar el efecto de SPARC en la progresión tumoral *in vivo*. Se injertó a ratones desnudos NIH (6 semanas de edad, Taconic Laboratories) después de la inyección subcutánea de 2×10^6 células en el flanco izquierdo. Se iniciaron los regímenes de tratamiento una vez que el tamaño medio tumoral era de 50-75 mm³ de tamaño. Las mediciones de los tumores se llevaron a cabo usando un calibre manual (Fisher) dos veces a la semana y las mediciones del peso se efectuaron de manera concurrente hasta terminar el estudio. La quimioterapia se administró usando un régimen de ciclos de 3 semanas durante un total de 6 ciclos: inyecciones intraperitoneales de 25 mg/kg de 5-FU o 25 mg/kg de CPT-11 tres veces en la semana 1 de cada ciclo, seguido de periodos sin tratamiento de 2 semanas. La pauta de dosificación para SPARC fue de 100 μ l de SPARC tres veces a la semana hasta terminar el ciclo de quimioterapia.

Ejemplo 2. Expresión de SPARC en células resistentes a la quimioterapia

Se usaron dos clones resistentes a la quimioterapia (MIP-4FUR y MIP-ETOR), apoyándose en los ensayos de formación de colonias (fig. 3) y el ensayo TUNEL (fig. 4) para la detección de SPARC en las células resistentes a la quimioterapia. El análisis de micromatriz identificó una serie de genes subexpresados en las células resistentes, incluyendo SPARC.

La subexpresión a nivel de expresión génica también se tradujo en menores niveles de proteína de SPARC en las líneas celulares resistentes a la quimioterapia (fig. 5A). Esta característica no fue única para las líneas celulares resistentes desarrolladas únicamente para los fines del presente estudio, ya que otra línea celular de sarcoma uterino bien establecida, MES-SA, también mostró una expresión disminuida de SPARC cuando es resistente a un agente quimioterapéutico diferente, doxorubicina (fig. 5B). Además, en muestras humanas de patología normales, el nivel de expresión de proteína de SPARC parece ser máximo en las vellosidades, con un gradiente descendiente hacia las criptas colónicas. Esta expresión variable se pierde en el caso de neoplasias malignas, con una disminución general en la expresión de SPARC en el adenocarcinoma de colon en diversos estadios.

La figura 19 muestra los niveles de ARNm y proteína de SPARC en líneas celulares de cáncer colorrectal sensibles y resistentes a la quimioterapia. (A) El diagrama de análisis de agrupación de micromatriz de oligonucleótidos (panel izquierdo) revela que la expresión génica de SPARC es significativamente menor en líneas celulares resistentes a la quimioterapia, lo que se confirmó mediante RT-PCR semi-cuantitativa (panel derecho). (B) La detección de los niveles de expresión de SPARC en una línea celular de sarcoma uterino emparejada sensible a la quimioterapia (MES-SA) y resistente a la doxorubicina (MES-SA/DX5) muestra una disminución similar en la expresión en la línea celular resistente. En líneas celulares de cáncer de mama, MDA435 tenía niveles de expresión ligeramente mayores de SPARC que MCF-7. Se detectaron bajos niveles de expresión en la línea celular de cáncer pancreático (CRL 1420), la línea celular de cáncer de pulmón (JMN 1B), las líneas celulares de cáncer colorrectal (RKO, CCL 227, HT 29). Se detectaron altos niveles de expresión de SPARC en la línea celular de colon normal (CRL 1541) y en una línea celular de cáncer de colon (HCT 116). (C) La expresión de proteína de SPARC verifica que hay una disminución significativa en esta proteína en los clones resistentes de MIP101 (líneas celulares resistentes: MIP/5FU, MIP/CPT, MIP/ETO, MIPT/CIS) en comparación con la línea celular progenitora normal (carril 5, MIP101). De manera similar, otro grupo de líneas celulares resistentes cuyo origen era sarcoma uterino (MES-SA/DX5, sarcoma uterino resistente a la doxorubicina) muestra una expresión de SPARC disminuida en comparación con las líneas celulares progenitoras sensibles (MES-SA, progenitora de sarcoma uterino).

La figura 20 muestra la expresión de proteína de SPARC en epitelio colónico humano. (A) El colon normal muestra un patrón diferencial de expresión de proteína de SPARC con mayores niveles de la proteína en las células superficiales próximas al lumen y un gradiente de expresión disminuida hacia las criptas. Los niveles de proteína de SPARC en (B) adenocarcinoma de colon, (C) adenocarcinoma mucinoso y (D) adenocarcinoma de colon

metastásico al hígado muestran bajos niveles de proteína de SPARC de manera difusa en el epitelio maligno. Secciones trasversales de 6 µm, aumento 20X.

Ejemplo 3. El polipéptido de SPARC sensibiliza a las células resistentes al tratamiento con 5-FU

Para delinear adicionalmente este papel potencial, se evaluó la respuesta de las células MIP101 resistentes (fig. 6) a SPARC exógeno para revertir el fenotipo resistente. Tal como se indicaba por los experimentos iniciales, LAS células MIP101 resistentes a 5-FU (MIP-5FUR) no podían activarse para sufrir apoptosis con 5-FU a una concentración de 500 µM, mientras que un número significativo de células de la línea celular progenitora sensible sufrió apoptosis después de exponerlas a una concentración similar de 5-FU. Se observó un hallazgo significativo con la exposición exógena con SPARC de células resistentes: la incubación de los clones resistentes con SPARC durante un periodo de 24 horas seguido de una exposición de 12 h a un agente quimioterapéutico fue suficiente para revertir el fenotipo resistente, ya que de nuevo se detectaron células apoptóticas mediante el ensayo TUNEL en células expuestas a concentraciones de quimioterapia que no estimularon previamente la muerte celular. La incubación solo con SPARC exógeno sin exposición posterior a la quimioterapia no indujo apoptosis ni en las células MIP101 progenitoras ni en las células resistentes.

La figura 21 muestra la evaluación del efecto de SPARC para influir en la sensibilidad de las células a la quimioterapia. (A) Efecto de la exposición de células MIP/5FU a SPARC exógeno combinado con 5-FU *in vitro*. La evaluación de la apoptosis mediante ensayo TUNEL muestra células teñidas positivamente en células MIP101 sensibles expuestas a 1000 µM de 5-FU (a, tinción TUNEL; b tinción DAPI) pero carentes de apoptosis en el fenotipo resistente (c, tinción TUNEL; d, tinción DAPI) después de exposición a una concentración similar de 5-FU. Sin embargo, después de una exposición de 24 h a SPARC (5 µg/ml), las células resistentes a 5-FU de nuevo fueron sensibles a 1000 µM de 5-FU según se muestra por las células teñidas positivamente en TUNEL (e; f, tinción DAPI), lo que indica la presencia de células apoptóticas. Esta es la primera indicación de que la exposición exógena a SPARC revierte el fenotipo resistente de las células resistentes a 5-FU y por lo tanto sugiere que SPARC puede funcionar como sensibilizante a la quimioterapia. (B) Las líneas celulares MIP101 transducidas de manera estable que sobreexpresan SPARC (MIP/SP) y de control (MIP/Zeo) expuestas a concentraciones crecientes de quimioterapia (5-FU, CTP-11 y etopósido) mostraron menos colonias celulares de células MIP/SP cuando se expusieron a concentraciones menores de fármaco que las células MIP/Zeo, indicando de este modo una mayor sensibilidad de los clones que sobreexpresan SPARC a la quimioterapia ya que sobrevivieron menos células a concentraciones relativamente más bajas de quimioterapia. (B) Un mayor número de células MIP101 que sobreexpresan SPARC (MIP/SP) sufren apoptosis después de una exposición de 12 h a varios agentes quimioterapéuticos (ETO = etopósido, CIS = cisplatino, 5-FU = 5-fluorouracilo, CPT = CPT-11) en comparación con las células de control (MIP/Zeo) ($p < 0,05$). El análisis de la apoptosis después del etiquetado con anexina V mediante citometría de flujo representa los resultados de tres estudios independientes efectuados por triplicado. El ensayo clonogénico (B) es un experimento representativo que se repitió tres veces con resultados similares.

Ejemplo 4. El polinucleótido de SPARC sensibiliza a células recombinantes a varios tratamientos quimioterapéuticos

Para probar esta hipótesis, se transfectaron células MIP101 con SPARC con el fin de generar sistemas de sobreexpresión para estudios *in vitro* adicionales. Se usaron dos clones que sobreexpresaban SPARC (clones 4, 5; fig 7) para estudios posteriores.

La sensibilidad de los transfectantes de SPARC a diversos agentes quimioterapéuticos se evaluó mediante ensayo de formación de colonias, que mostró que los clones que sobreexpresaban SPARC eran incapaces de formar colonias tumorigénicas a mayores concentraciones de quimioterapia en comparación con las líneas celulares progenitoras. De manera similar, el análisis FACS de las poblaciones celulares inducidas para sufrir apoptosis después de la exposición a agentes quimioterapéuticos mostró un desplazamiento drástico hacia la apoptosis temprana en los transfectantes de SPARC (fig. 8D) después de una exposición de 12 h a la quimioterapia. Una población menor de células de la línea celular progenitora sufrió apoptosis después de la inducción solo con quimioterapia (fig. 8C). En su conjunto, parecía haber un aumento de al menos 2 veces en la población de células que sobreexpresan SPARC que sufren apoptosis en comparación con la línea celular MIP101 progenitora después de la exposición a diversos agentes quimioterapéuticos (fig. 9). La figura 10 muestra la respuesta de los transfectantes de SPARC a agentes quimioterapéuticos.

Ejemplo 5. Se observa sensibilización de SPARC *in vivo*

La sensibilidad aumentada a la quimioterapia *in vitro* se tradujo al sistema de modelo *in vivo*, mostrando dos de cuatro animales regresión tumoral completa en animales trasplantados con transfectantes de SPARC después de 6 ciclos de quimioterapia (fig. 11). El resto de animales injertados con transfectantes de SPARC tuvieron una reducción drástica en la velocidad de crecimiento tumoral en comparación con los animales injertados con MIP101 progenitoras. Todos los animales de control (xenoinjertos de MIP101 tratados con quimioterapia) tenían tumores $>400 \text{ mm}^2$ a los 50 días después de iniciar el tratamiento quimioterapéutico, mientras que los animales injertados con transfectante de SPARC que no sufrieron regresión tumoral completa, tenían tumores que seguían siendo de $<300 \text{ mm}^2$ a los 140 días después del inicio de la quimioterapia (resultado no mostrado).

Ejemplo 6: Método para seleccionar un agente que modula la expresión de un polipéptido de SPARC

La exploración de un modulador de la expresión del polipéptido de SPARC puede llevarse a cabo como una sencilla exploración basada en células de mamífero. En primer lugar se preincuba una línea celular de cultivo tisular de mamífero, por ejemplo, células HeLa con moléculas pequeñas candidatas aleatorias. Entonces se exploran los clones celulares usando transferencias de Western anti-SPARC o ELISA. Como alternativa, se lleva a cabo una reacción RT-PCR para examinar la modulación de la expresión de ARNm de SPARC.

Ejemplo 7: Modelo de terapia en animales adicional

Se llevó a cabo un modelo de terapia en varios animales y los resultados se muestran en la figura 14.

En la figura 14, los animales xenoinjertados a los que se injertan MIP101 o MIP/SP tratados con diferentes agentes quimioterapéuticos (5-FU o CPT-11) muestran una velocidad más rápida de regresión tumoral de los xenoinjertos de tumor de MIP/SP en comparación con xenoinjertos de tumor de la línea celular progenitora MIP101 (MIP). Dos de cuatro animales que portaban xenoinjertos de MIP/SP experimentaron regresión tumoral completa, mientras que los otros dos tuvieron una velocidad de crecimiento del tumor mucho más lenta en comparación con los animales de control que portaban MIP101 expuestos a un régimen de tratamiento similar. Los animales representativos con xenoinjertos de tumor de MIP-SPARC tratados con 2 ciclos de 5-FU tuvieron remisión completa a los 23 días después del trasplante o tumores significativamente más pequeños después de solo 2 ciclos de CPT-11 en comparación con un animal trasplantado con un xenoinjerto de MIP101 progenitor.

En la figura 15, más animales con xenoinjertos de células MIP/SP mostraron evidencias de regresión tumoral completa más temprana en el periodo posterior al tratamiento con radiación que los animales con xenoinjertos de células MIP/Zeo de control. A las 15 semanas después de la radioterapia, ninguno de los animales con xenoinjerto de MIP/SP tenían evidencias de tumores, mientras que el 30% de los xenoinjertados con MIP/Zeo seguían portando tumores (n = 10 animales/grupo; dosis total de radiación: 100Gy).

En la figura 16, el tratamiento combinado con SPARC (IP, intraperitoneal) y 5-FU dio como resultado una regresión tumoral que fue significativamente mayor que el tratamiento solo con 5-FU a los 51 días después del inicio del tratamiento- (B) Este tratamiento combinado de SPARC (IP) y 5-FU dio como resultado regresión tumoral completa en varios animales a los 84 días después del tratamiento, mientras que esto no se observó en animales tratados solo con 5-FU. (Media $\pm$ DT, n = 6 animales/grupo).

En la figura 17, el tratamiento combinado con SPARC (SC, subcutáneo) y 5-FU dio como resultado una regresión tumoral que fue significativamente mayor que el tratamiento solo con 5-FU durante el periodo de tratamiento. Este tratamiento combinado de SPARC (SC) y 5-FU dio como resultado regresión tumoral completa en varios animales a los 42 días después del tratamiento, mientras que esto no se observó en animales tratados solo con 5-FU. (Media $\pm$ DT, n = 6 animales/grupo).

En la figura 18, los animales injertados con células resistentes MIP/5FU tratados bien con 5-FU solo o con combinación de SPARC y 5-FU mostraron que el rápido crecimiento del tumor continuó en los animales tratados solo con 5-FU, mientras que se observó una dramática regresión tumoral en los animales tratados con la terapia combinada comenzando a los 28 días después del tratamiento. Varios animales que recibieron terapia combinada de SPARC y 5-FU mostraron regresión tumoral completa a los 117 días después del tratamiento, (Media $\pm$ DT, n = 6 animales/grupo).

Listado de secuencias

- <110> Dana Farber Cancer Institute
- <120> Sensibilizador a la terapia para el cáncer
- <130> P023631EPA
- <140> EP 12183770.2
- <141> 14-01-2004
- <150> EP 04702173.8
- <151> 14-01-2004
- <150> US 60/440.009
- <151> 14-01-2003
- <160> 81

<170> PatentIn versión 3.5

<210> 1

<211> 303

<212> PRT

<213> *Homo sapiens*

<400> 1

```

Met Arg Ala Trp Ile Phe Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu
1          5          10          15

Ala Ala Pro Gln Gln Glu Ala Leu Pro Asp Glu Thr Glu Val Val Glu
20          25          30

Glu Thr Val Ala Glu Val Thr Glu Val Ser Val Gly Ala Asn Pro Val
35          40          45

Gln Val Glu Val Gly Glu Phe Asp Asp Gly Ala Glu Glu Thr Glu Glu
50          55          60

Glu Val Val Ala Glu Asn Pro Cys Gln Asn His His Cys Lys His Gly
65          70          75          80

Lys Val Cys Glu Leu Asp Glu Asn Asn Thr Pro Met Cys Val Cys Gln
85          90          95

Asp Pro Thr Ser Cys Pro Ala Pro Ile Gly Glu Phe Glu Lys Val Cys
100         105         110

Ser Asn Asp Asn Lys Thr Phe Asp Ser Ser Cys His Phe Phe Ala Thr
115         120         125

Lys Cys Thr Leu Glu Gly Thr Lys Lys Gly His Lys Leu His Leu Asp
130         135         140

```

ES 2 565 496 T3

Tyr Ile Gly Pro Cys Lys Tyr Ile Pro Pro Cys Leu Asp Ser Glu Leu
 145 150 155 160
 Thr Glu Phe Pro Leu Arg Met Arg Asp Trp Leu Lys Asn Val Leu Val
 165 170 175
 Thr Leu Tyr Glu Arg Asp Glu Asp Asn Asn Leu Leu Thr Glu Lys Gln
 180 185 190
 Lys Leu Arg Val Lys Lys Ile His Glu Asn Glu Lys Arg Leu Glu Ala
 195 200 205
 Gly Asp His Pro Val Glu Leu Leu Ala Arg Asp Phe Glu Lys Asn Tyr
 210 215 220
 Asn Met Tyr Ile Phe Pro Val His Trp Gln Phe Gly Gln Leu Asp Gln
 225 230 235 240
 His Pro Ile Asp Gly Tyr Leu Ser His Thr Glu Leu Ala Pro Leu Arg
 245 250 255
 Ala Pro Leu Ile Pro Met Glu His Cys Thr Thr Arg Phe Phe Glu Thr
 260 265 270
 Cys Asp Leu Asp Asn Asp Lys Tyr Ile Ala Leu Asp Glu Trp Ala Gly
 275 280 285
 Cys Phe Gly Ile Lys Gln Lys Asp Ile Asp Lys Asp Leu Val Ile
 290 295 300

<210> 2
 <211> 302
 <212> PRT
 <213> *Mus musculus*

5

<400> 2

Met Arg Ala Trp Ile Phe Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu
 1 5 10 15
 Ala Ala Pro Gln Gln Thr Glu Val Ala Glu Glu Ile Val Glu Glu Glu
 20 25 30
 Thr Val Val Glu Glu Thr Gly Val Pro Val Gly Ala Asn Pro Val Gln
 35 40 45

ES 2 565 496 T3

Val Glu Met Gly Glu Phe Glu Asp Gly Ala Glu Glu Thr Val Glu Glu
50 55 60

Val Val Ala Asp Asn Pro Cys Gln Asn His His Cys Lys His Gly Lys
65 70 75 80

Val Cys Glu Leu Asp Glu Ser Asn Thr Pro Met Cys Val Cys Gln Asp
85 90 95

Pro Thr Ser Cys Pro Ala Pro Ile Gly Glu Phe Glu Lys Val Cys Ser
100 105 110

Asn Asp Asn Lys Thr Phe Asp Ser Ser Cys His Phe Phe Ala Thr Lys
115 120 125

Cys Thr Leu Glu Gly Thr Lys Lys Gly His Lys Leu His Leu Asp Tyr
130 135 140

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

Glu Phe Pro Leu Arg Met Arg Asp Trp Leu Lys Asn Val Leu Val Thr
165 170 175

Leu Tyr Glu Arg Asp Glu Gly Asn Asn Leu Leu Thr Glu Lys Gln Lys
180 185 190

Leu Arg Val Lys Lys Ile His Glu Asn Glu Lys Arg Leu Glu Ala Gly
195 200 205

Asp His Pro Val Glu Leu Leu Ala Arg Asp Phe Glu Lys Asn Tyr Asn
210 215 220

Met Tyr Ile Phe Pro Val His Trp Gln Phe Gly Gln Leu Asp Gln His
225 230 235 240

Pro Ile Asp Gly Tyr Leu Ser His Thr Glu Leu Ala Pro Leu Arg Ala
245 250 255

Pro Leu Ile Pro Met Glu His Cys Thr Thr Arg Phe Phe Glu Thr Cys
260 265 270

Asp Leu Asp Asn Asp Lys Tyr Ile Ala Leu Glu Glu Trp Ala Gly Cys
275 280 285

Phe Gly Ile Lys Glu Gln Asp Ile Asn Lys Asp Leu Val Ile
290 295 300

<210> 3
<211> 345
<212> PRT
<213> *Homo sapiens*

<400> 3

```

Met Leu Pro Cys Phe Gln Leu Leu Arg Ile Gly Gly Gly Arg Gly Gly
1           5           10           15

Asp Leu Tyr Thr Phe His Pro Pro Ala Gly Ala Gly Cys Thr Tyr Arg
20           25           30

Leu Gly His Arg Ala Asp Leu Cys Asp Val Ala Leu Arg Pro Gln Gln
35           40           45

Glu Pro Gly Leu Ile Ser Gly Ile His Ala Glu Leu His Ala Glu Pro
50           55           60

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

Thr Leu Val Asn Asn Val Arg Leu Pro Arg Gly His Arg Leu Glu Leu
85           90           95

Ser Asp Gly Asp Leu Leu Thr Phe Gly Pro Glu Gly Pro Pro Gly Thr
100          105          110

Ser Pro Ser Glu Phe Tyr Phe Met Phe Gln Gln Val Arg Val Lys Pro
115          120          125

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

Val Gly Ala Gly Phe Arg Pro Met Leu Pro Ser Gln Gly Ala Pro Gln
145          150          155          160

Arg Pro Leu Ser Thr Phe Ser Pro Ala Pro Lys Ala Thr Leu Ile Leu
165          170          175

Asn Ser Ile Gly Ser Leu Ser Lys Leu Arg Pro Gln Pro Leu Thr Phe
180          185          190

Ser Pro Ser Trp Gly Gly Pro Lys Ser Leu Pro Val Pro Ala Pro Pro
195          200          205

Gly Glu Val Gly Thr Thr Pro Ser Ala Pro Pro Gln Arg Asn Arg Arg
210          215          220

Lys Ser Val His Arg Val Leu Ala Glu Leu Asp Asp Glu Ser Glu Pro
225          230          235          240

Pro Glu Asn Pro Pro Pro Val Leu Met Glu Pro Arg Lys Lys Leu Arg
245          250          255

Val Asp Lys Ala Pro Leu Thr Pro Thr Gly Asn Arg Arg Gly Arg Pro
260          265          270

Arg Lys Tyr Pro Val Ser Ala Pro Met Ala Pro Pro Ala Val Gly Gly
275          280          285

Gly Glu Pro Cys Ala Ala Pro Cys Cys Cys Leu Pro Gln Glu Glu Thr
290          295          300

Val Ala Trp Val Gln Cys Asp Gly Cys Asp Val Trp Phe His Val Ala
305          310          315          320

Cys Val Gly Cys Ser Ile Gln Ala Ala Arg Glu Ala Asp Phe Arg Cys
325          330          335

Pro Gly Cys Arg Ala Gly Ile Gln Thr
340          345

```

<210> 4
 <211> 345
 <212> PRT
 <213> *Homo sapiens*

 <400> 4


```

Met Leu Pro Cys Phe Gln Leu Leu Arg Ile Gly Gly Gly Arg Gly Gly
1      5      10      15
Asp Leu Tyr Thr Phe His Pro Pro Ala Gly Ala Gly Cys Thr Tyr Arg
20      25      30
Leu Gly His Arg Ala Asp Leu Cys Asp Val Ala Leu Arg Pro Gln Gln
35      40      45
Glu Pro Gly Leu Ile Ser Gly Ile His Ala Glu Leu His Ala Glu Pro
50      55      60
Arg Gly Asp Asp Trp Arg Val Ser Leu Glu Asp His Ser Ser Gln Gly
65      70      75      80
Thr Leu Val Asn Asn Val Arg Leu Pro Arg Gly His Arg Leu Glu Leu
85      90      95
Ser Asp Gly Asp Leu Leu Thr Phe Gly Pro Glu Gly Pro Pro Gly Thr
100     105     110
Ser Pro Ser Glu Phe Tyr Phe Met Phe Gln Gln Val Arg Val Lys Pro
115     120     125
Gln Asp Phe Ala Ala Ile Thr Ile Pro Arg Ser Arg Gly Glu Ala Arg
130     135     140
Val Gly Ala Gly Phe Arg Pro Met Leu Pro Ser Gln Gly Ala Pro Gln
145     150     155     160
Arg Pro Leu Ser Thr Phe Ser Pro Ala Pro Lys Ala Thr Leu Ile Leu
165     170     175
Asn Ser Ile Gly Ser Leu Ser Lys Leu Arg Pro Gln Pro Leu Thr Phe
180     185     190
Ser Pro Ser Trp Gly Gly Pro Lys Ser Leu Pro Val Pro Ala Pro Pro
195     200     205
Gly Glu Val Gly Thr Thr Pro Ser Ala Pro Pro Gln Arg Asn Arg Arg
210     215     220
Lys Ser Val His Arg Val Leu Ala Glu Leu Asp Asp Glu Ser Glu Pro
225     230     235     240
Leu Glu Asn Pro Pro Pro Val Leu Met Glu Pro Arg Lys Lys Leu Arg
245     250     255
Val Asp Lys Ala Pro Leu Thr Pro Thr Gly Asn Arg Arg Gly Arg Pro
260     265     270
Arg Lys Tyr Pro Val Ser Ala Pro Met Ala Pro Pro Ala Val Gly Gly
275     280     285
Gly Glu Pro Cys Ala Ala Pro Cys Cys Cys Leu Pro Gln Glu Glu Thr
290     295     300
Val Ala Trp Val Gln Cys Asp Gly Cys Asp Val Trp Phe His Val Ala
305     310     315     320
Cys Val Gly Cys Ser Ile Gln Ala Ala Arg Glu Ala Asp Phe Arg Cys
325     330     335
Pro Gly Cys Arg Ala Gly Ile Gln Thr
340     345

```

<210> 5
 <211> 359
 <212> PRT
 <213> *Homo sapiens*
 <400> 5

Met Leu Pro Cys Phe Gln Leu Leu Arg Ile Gly Gly Gly Arg Gly Gly
1 5 10 15

Asp Leu Tyr Thr Phe His Pro Pro Ala Gly Val Gly Cys Thr Tyr Arg
20 25 30

Leu Gly His Arg Ala Asp Leu Cys Asp Val Ala Leu Arg Pro Gln Gln
35 40 45

Glu Pro Gly Leu Ile Ser Gly Ile His Ala Glu Leu His Ala Glu Pro
50 55 60

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

Thr Leu Val Asn Asn Val Arg Leu Pro Arg Gly His Arg Leu Glu Leu
85 90 95

Ser Asp Gly Asp Leu Leu Thr Phe Gly Pro Glu Gly Pro Pro Gly Thr
100 105 110

Ser Pro Ser Glu Phe Tyr Phe Met Phe Gln Gln Val Arg Val Lys Pro
115 120 125

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

Val Gly Ala Gly Phe Arg Pro Met Leu Pro Ser Gln Gly Ala Pro Gln
145 150 155 160

Arg Pro Leu Ser Thr Phe Ser Pro Ala Pro Lys Ala Thr Leu Ile Leu
165 170 175

Asn Ser Ile Gly Ser Leu Ser Lys Leu Arg Pro Gln Pro Leu Thr Phe
180 185 190

Ser Pro Ser Trp Gly Gly Pro Lys Ser Leu Pro Val Pro Ala Pro Pro
195 200 205

Gly Glu Val Gly Thr Thr Pro Ser Ala Pro Pro Gln Arg Asn Arg Arg
210 215 220

Lys Ser Val His Arg Val Leu Ala Glu Leu Asp Asp Glu Ser Glu Pro
225 230 235 240

Pro Glu Asn Pro Pro Pro Val Leu Met Glu Pro Arg Lys Lys Leu Arg
245 250 255

Val Asp Lys Ala Pro Leu Thr Pro Thr Gly Asn Ala Arg Gly Arg Pro
260 265 270

Arg Lys Tyr Pro Val Ser Ala Pro Met Ala Pro Pro Ala Val Gly Ala
275 280 285

Gly Ser Pro Val Gln Leu Leu Val Ala Ala Cys Pro Arg Lys Arg Gln
290 295 300

Trp Pro Gly Phe Ser Val Met Ala Val Thr Ser Gly Ser Met Trp Pro
305 310 315 320

Val Leu Ala Ala Ala Ser Arg Leu Pro Gly Arg Pro Thr Ser Asp Ala
325 330 335

Gln Gly Ala Gly Leu Ala Phe Ser Leu Arg Ser Thr Ala Lys Ala Pro
340 345 350

Ser Asp Thr Pro Ala His Glu
355

<210> 6
<211> 345
<212> PRT
<213> *Homo sapiens*

<400> 6

```

Met Leu Pro Cys Phe Gln Leu Leu Arg Ile Gly Gly Gly Arg Gly Gly
 1           5           10           15

Asp Leu Tyr Thr Phe His Pro Pro Ala Gly Ala Gly Cys Thr Tyr Arg
 20           25           30

Leu Gly His Arg Ala Asp Leu Cys Asp Val Ala Leu Arg Pro Gln Gln
 35           40           45

Glu Pro Gly Leu Ile Ser Gly Ile His Ala Glu Leu His Ala Glu Pro
 50           55           60

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

Thr Leu Val Asn Asn Val Arg Leu Pro Arg Gly His Arg Leu Glu Leu
 85           90           95

Ser Asp Gly Asp Leu Leu Thr Phe Gly Pro Glu Gly Pro Pro Gly Thr
 100           105           110

Ser Pro Ser Glu Phe Tyr Phe Met Phe Gln Gln Val Arg Val Lys Pro
 115           120           125

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

Val Gly Ala Gly Phe Arg Pro Met Leu Pro Ser Gln Gly Ala Pro Gln
 145           150           155           160

Arg Pro Leu Ser Thr Phe Ser Pro Ala Pro Lys Ala Thr Leu Ile Leu
 165           170           175

Asn Ser Ile Gly Ser Leu Ser Lys Leu Arg Pro Gln Pro Leu Thr Phe
 180           185           190

Ser Pro Ser Trp Gly Gly Pro Lys Ser Leu Pro Val Pro Ala Pro Pro
 195           200           205

Gly Glu Val Gly Thr Thr Pro Ser Ala Pro Pro Gln Arg Asn Arg Arg
 210           215           220

Lys Ser Val His Arg Val Leu Ala Glu Leu Asp Asp Glu Ser Glu Pro
 225           230           235           240

Pro Glu Asn Pro Pro Pro Val Leu Met Glu Pro Arg Lys Lys Leu Arg
 245           250           255

Val Asp Lys Ala Pro Leu Thr Pro Thr Gly Asn Arg Arg Gly Arg Pro
 260           265           270

Arg Lys Tyr Pro Val Ser Ala Pro Met Ala Pro Pro Ala Val Gly Gly
 275           280           285

Gly Glu Pro Cys Ala Ala Pro Cys Cys Cys Leu Pro Gln Glu Glu Thr
 290           295           300

Val Ala Trp Val Gln Cys Asp Gly Cys Asp Val Trp Phe His Val Ala
 305           310           315           320

Cys Val Gly Cys Ser Ile Gln Ala Ala Arg Glu Ala Asp Phe Arg Cys
 325           330           335

Pro Gly Cys Arg Ala Gly Ile Gln Thr
 340           345

```

5

<210> 7
 <211> 359
 <212> PRT
 <213> *Homo sapiens*

<400> 7

```

Met Leu Pro Cys Phe Gln Leu Leu Arg Ile Gly Gly Gly Arg Gly Gly
1      5      10      15

Asp Leu Tyr Thr Phe His Pro Pro Ala Gly Val Gly Cys Thr Tyr Arg
      20      25      30

Leu Gly His Arg Ala Asp Leu Cys Asp Val Ala Leu Arg Pro Gln Gln
      35      40      45
Glu Pro Gly Leu Ile Ser Gly Ile His Ala Glu Leu His Ala Glu Pro
      50      55      60

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

Thr Leu Val Asn Asn Val Arg Leu Pro Arg Gly His Arg Leu Glu Leu
      85      90      95

Ser Asp Gly Asp Leu Leu Thr Phe Gly Pro Glu Gly Pro Pro Gly Thr
      100      105      110

Ser Pro Ser Glu Phe Tyr Phe Met Phe Gln Gln Val Arg Val Lys Pro
      115      120      125

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

Val Gly Ala Gly Phe Arg Pro Met Leu Pro Ser Gln Gly Ala Pro Gln
145      150      155      160

Arg Pro Leu Ser Thr Phe Ser Pro Ala Pro Lys Ala Thr Leu Ile Leu
      165      170      175

Asn Ser Ile Gly Ser Leu Ser Lys Leu Arg Pro Gln Pro Leu Thr Phe
      180      185      190

Ser Pro Ser Trp Gly Gly Pro Lys Ser Leu Pro Val Pro Ala Pro Pro
      195      200      205

Gly Glu Val Gly Thr Thr Pro Ser Ala Pro Pro Gln Arg Asn Arg Arg
210      215      220

Lys Ser Val His Arg Val Leu Ala Glu Leu Asp Asp Glu Ser Glu Pro
225      230      235      240

Pro Glu Asn Pro Pro Pro Val Leu Met Glu Pro Arg Lys Lys Leu Arg
      245      250      255

Val Asp Lys Ala Pro Leu Thr Pro Thr Gly Asn Ala Arg Gly Arg Pro
      260      265      270

Arg Lys Tyr Pro Val Ser Ala Pro Met Ala Pro Pro Ala Val Gly Ala
      275      280      285

Gly Ser Pro Val Gln Leu Leu Val Ala Ala Cys Pro Arg Lys Arg Gln
290      295      300

Trp Pro Gly Phe Ser Val Met Ala Val Thr Ser Gly Ser Met Trp Pro
305      310      315      320

Val Leu Ala Ala Ala Ser Arg Leu Pro Gly Arg Pro Thr Ser Asp Ala
      325      330      335

Gln Gly Ala Gly Leu Ala Phe Ser Leu Arg Ser Thr Ala Lys Ala Pro
      340      345      350

Ser Asp Thr Pro Ala His Glu
      355

```

5

<210> 8
<211> 650
<212> PRT

<213> *Mus musculus*

<400> 8

```

Met Lys Ala Val Leu Leu Leu Cys Ala Leu Gly Thr Ala Val Ala
 1           5           10           15

Ile Pro Thr Ser Thr Arg Phe Leu Phe Asp His Ser Asn Pro Thr Thr
          20           25           30

Ala Thr Leu Val Thr Pro Glu Asp Ala Thr Val Pro Ile Ala Gly Val
          35           40           45

Glu Ala Thr Ala Asp Ile Glu Asn His Pro Asn Asp Lys Ala Glu Lys
 50           55           60

Pro Ser Ala Leu Asn Ser Glu Glu Glu Thr His Glu Gln Ser Thr Glu
65           70           75           80

Gln Asp Lys Thr Tyr Ser Phe Glu Val Asp Leu Lys Asp Glu Glu Asp
          85           90           95

Gly Asp Gly Asp Leu Ser Val Asp Pro Thr Glu Gly Thr Leu Thr Leu
          100          105          110

Asp Leu Gln Glu Gly Thr Ser Glu Pro Gln Gln Lys Ser Leu Pro Glu
          115          120          125

Asn Gly Asp Phe Pro Ala Thr Val Ser Thr Ser Tyr Val Asp Pro Asn
130          135          140

Gln Arg Ala Asn Ile Thr Lys Gly Lys Glu Ser Gln Glu Gln Pro Val
145          150          155          160

Ser Asp Ser His Gln Gln Pro Asn Glu Ser Ser Lys Gln Thr Gln Asp
          165          170          175

Leu Lys Ala Glu Glu Ser Gln Thr Gln Asp Pro Asp Ile Pro Asn Glu
          180          185          190

Glu Glu Glu Glu Glu Glu Asp Glu Glu Glu Glu Glu Glu Glu Pro
          195          200          205

Glu Asp Ile Gly Ala Pro Ser Asp Asn Gln Glu Glu Gly Lys Glu Pro
210          215          220

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

Asp Asp Thr Leu Glu Glu Ser Ser Gln Pro Thr Gln Ile Ser Lys Thr
          245          250          255

Glu Lys His Gln Ser Glu Gln Gly Asn Gln Gly Gln Glu Ser Asp Ser
          260          265          270

Glu Ala Glu Gly Glu Asp Lys Ala Ala Gly Ser Lys Glu His Ile Pro
          275          280          285

His Thr Glu Gln Gln Asp Gln Glu Gly Lys Ala Gly Leu Glu Ala Ile
290          295          300

Gly Asn Gln Lys Asp Thr Asp Glu Lys Ala Val Ser Thr Glu Pro Thr

```

```

305              310              315              320
Asp Ala Ala Val Val Pro Arg Ser His Gly Gly Ala Gly Asp Asn Gly
          325          330          335
Gly Gly Asp Asp Ser Lys His Gly Ala Gly Asp Asp Tyr Phe Ile Pro
          340          345          350
Ser Gln Glu Phe Leu Glu Ala Glu Arg Met His Ser Leu Ser Tyr Tyr
          355          360          365
Leu Lys Tyr Gly Gly Gly Glu Glu Thr Thr Thr Gly Glu Ser Glu Asn
          370          375          380
Arg Arg Glu Ala Ala Asp Asn Gln Glu Ala Lys Lys Ala Glu Ser Ser
385          390          395          400
Pro Asn Ala Glu Pro Ser Asp Glu Gly Asn Ser Arg Glu His Ser Ala
          405          410          415
Gly Ser Cys Thr Asn Phe Gln Cys Lys Arg Gly His Ile Cys Lys Thr
          420          425          430
Asp Pro Gln Gly Lys Pro His Cys Val Cys Gln Asp Pro Glu Thr Cys
          435          440          445
Pro Pro Ala Lys Ile Leu Asp Gln Ala Cys Gly Thr Asp Asn Gln Thr
          450          455          460
Tyr Ala Ser Ser Cys His Leu Phe Ala Thr Lys Cys Arg Leu Glu Gly
465          470          475          480
Thr Lys Lys Gly His Gln Leu Gln Leu Asp Tyr Phe Gly Ala Cys Lys
          485          490          495
Ser Ile Pro Ala Cys Thr Asp Phe Glu Val Ala Gln Phe Pro Leu Arg
          500          505          510
Met Arg Asp Trp Leu Lys Asn Ile Leu Met Gln Leu Tyr Glu Pro Asn
          515          520          525
Pro Lys His Gly Gly Tyr Leu Asn Glu Lys Gln Arg Ser Lys Val Lys
          530          535          540
Lys Ile Tyr Leu Asp Glu Lys Arg Leu Leu Ala Gly Asp His Pro Ile
545          550          555          560
Glu Leu Leu Leu Arg Asp Phe Lys Lys Asn Tyr His Met Tyr Val Tyr
          565          570          575
Pro Val His Trp Gln Phe Asn Glu Leu Asp Gln His Pro Ala Asp Arg
          580          585          590
Ile Leu Thr His Ser Glu Leu Ala Pro Leu Arg Ala Ser Leu Val Pro
          595          600          605
Met Glu His Cys Ile Thr Arg Phe Phe Glu Glu Cys Asp Pro Asn Lys
          610          615          620
Asp Lys His Ile Thr Leu Lys Glu Trp Gly His Cys Phe Gly Ile Lys
625          630          635          640
Glu Glu Asp Ile Asp Glu Asn Leu Leu Phe
          645          650

```

<210> 9
 <211> 650
 <212> PRT
 <213> *Mus musculus*

<400> 9

Met	Lys	Ala	Val	Leu	Leu	Leu	Leu	Cys	Ala	Leu	Gly	Thr	Ala	Val	Ala		
1				5					10					15			
Ile	Pro	Thr	Ser	Thr	Arg	Phe	Leu	Ser	Asp	His	Ser	Asn	Pro	Thr	Thr		
			20					25					30				
Ala	Thr	Leu	Val	Thr	Pro	Glu	Asp	Ala	Thr	Val	Pro	Ile	Ala	Gly	Val		
		35				40					45						
Glu	Ala	Thr	Ala	Asp	Ile	Glu	Asn	His	Pro	Asn	Asp	Lys	Ala	Glu	Lys		
	50					55					60						
Pro	Ser	Ala	Leu	Asn	Ser	Glu	Glu	Glu	Thr	His	Glu	Gln	Ser	Thr	Glu		
65				70						75					80		
Gln	Asp	Lys	Thr	Tyr	Ser	Phe	Glu	Val	Asp	Leu	Lys	Asp	Glu	Glu	Asp		
				85					90					95			
Gly	Asp	Gly	Asp	Leu	Ser	Val	Asp	Pro	Thr	Glu	Gly	Thr	Leu	Thr	Leu		
			100					105					110				
Asp	Leu	Gln	Glu	Gly	Thr	Ser	Glu	Pro	Gln	Gln	Lys	Ser	Leu	Pro	Glu		
	115						120					125					
Asn	Gly	Asp	Phe	Pro	Ala	Thr	Val	Ser	Thr	Ser	Tyr	Val	Asp	Pro	Asn		
	130					135					140						
Gln	Arg	Ala	Asn	Ile	Thr	Lys	Gly	Lys	Glu	Ser	Gln	Glu	Gln	Pro	Val		
145				150						155					160		
Ser	Asp	Ser	His	Gln	Gln	Pro	Asn	Glu	Ser	Ser	Lys	Gln	Thr	Gln	Asp		
				165					170					175			
Leu	Lys	Ala	Glu	Glu	Ser	Gln	Thr	Gln	Asp	Pro	Asp	Ile	Pro	Asn	Glu		
		180						185					190				
Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Pro		
		195					200						205				
Glu	Asp	Ile	Gly	Ala	Pro	Ser	Asp	Asn	Gln	Glu	Glu	Gly	Lys	Glu	Pro		
	210					215					220						
Leu	Glu	Glu	Gln	Pro	Thr	Ser	Lys	Trp	Glu	Gly	Asn	Arg	Glu	Gln	Ser		
225				230						235					240		
Asp	Asp	Thr	Leu	Glu	Glu	Ser	Ser	Gln	Pro	Thr	Gln	Ile	Ser	Lys	Thr		
			245						250					255			
Glu	Lys	His	Gln	Ser	Glu	Gln	Gly	Asn	Gln	Gly	Gln	Glu	Ser	Asp	Ser		
		260						265					270				
Glu	Ala	Glu	Gly	Glu	Asp	Lys	Ala	Ser	Gly	Ser	Lys	Glu	His	Ile	Pro		
	275					280						285					

```

His Thr Glu Gln Gln Asp Gln Glu Gly Lys Ala Gly Leu Glu Ala Ile
 290                               295                 300

Gly Asn Gln Lys Asp Thr Asp Glu Lys Ala Val Ser Thr Glu Pro Thr
305                               310                 315                 320

Asp Ala Ala Val Val Pro Arg Ser His Gly Gly Ala Gly Asp Asn Gly
                               325                 330                 335

Gly Gly Asp Asp Ser Lys His Gly Ala Gly Asp Asp Tyr Phe Ile Pro
                               340                 345                 350

Ser Gln Glu Phe Leu Glu Ala Glu Arg Met His Ser Leu Ser Tyr Tyr
                               355                 360                 365

Leu Lys Tyr Gly Gly Gly Glu Glu Thr Thr Thr Gly Glu Ser Glu Asn
370                               375                 380

Gln Arg Glu Ala Ala Asp Asn Gln Glu Ala Lys Lys Ala Glu Ser Ser
385                               390                 395                 400

Pro Asn Ala Glu Pro Ser Asp Glu Gly Asn Ser Arg Glu His Ser Ala
                               405                 410                 415

Gly Ser Cys Thr Asn Phe Gln Cys Lys Arg Gly His Ile Cys Lys Thr
420                               425                 430

Asp Pro Gln Gly Lys Pro His Cys Val Cys Gln Asp Pro Glu Thr Cys
435                               440                 445

Pro Pro Ala Lys Ile Leu Asp Gln Ala Cys Gly Thr Asp Asn Gln Thr
450                               455                 460

Tyr Ala Ser Ser Cys His Leu Phe Ala Thr Lys Cys Arg Leu Glu Gly
465                               470                 475                 480

Thr Lys Lys Gly His Gln Leu Gln Leu Asp Tyr Phe Gly Ala Cys Lys
485                               490                 495

Ser Ile Pro Ala Cys Thr Asp Phe Glu Val Ala Gln Phe Pro Leu Arg
500                               505                 510

Met Arg Asp Trp Leu Lys Asn Ile Leu Met Gln Leu Tyr Glu Pro Asn
515                               520                 525

Pro Lys His Gly Gly Tyr Leu Asn Glu Lys Gln Arg Ser Lys Val Lys
530                               535                 540

Lys Ile Tyr Leu Asp Glu Lys Arg Leu Leu Ala Gly Asp His Pro Ile
545                               550                 555                 560

Glu Leu Leu Leu Arg Asp Phe Lys Lys Asn Tyr His Met Tyr Val Tyr
565                               570                 575

Pro Val His Trp Gln Phe Asn Glu Leu Asp Gln His Pro Ala Asp Arg
580                               585                 590

Ile Leu Thr His Ser Glu Leu Ala Pro Leu Arg Ala Ser Leu Val Pro
595                               600                 605

Met Glu His Cys Ile Thr Arg Phe Phe Glu Glu Cys Asp Pro Asn Lys
610                               615                 620

Asp Lys His Ile Thr Leu Lys Glu Trp Gly His Cys Phe Gly Ile Lys
625                               630                 635                 640

Glu Glu Asp Ile Asp Glu Asn Leu Leu Phe
645                               650

```

5
 <210> 10
 <211> 355
 <212> PRT
 <213> *Homo sapiens*

<400> 10

```

Met Ala Ala Pro Ala Phe Glu Pro Gly Arg Gln Ser Asp Leu Leu Val
1      5      10      15
Lys Leu Asn Arg Leu Met Glu Arg Cys Leu Arg Asn Ser Lys Cys Ile
20      25      30
Asp Thr Glu Ser Leu Cys Val Val Ala Gly Glu Lys Val Trp Gln Ile
35      40      45
Arg Val Asp Leu His Leu Leu Asn His Asp Gly Asn Ile Ile Asp Ala
50      55      60
Ala Ser Ile Ala Ala Ile Val Ala Leu Cys His Phe Arg Arg Pro Asp
65      70      75      80
Val Ser Val Gln Gly Asp Glu Val Thr Leu Tyr Thr Pro Glu Glu Arg
85      90      95
Asp Pro Val Pro Leu Ser Ile His His Met Pro Ile Cys Val Ser Phe
100     105     110
Ala Phe Phe Gln Gln Gly Thr Tyr Leu Leu Val Asp Pro Asn Glu Arg
115     120     125
Glu Glu Arg Val Met Asp Gly Leu Leu Val Ile Ala Met Asn Lys His
130     135     140
Arg Glu Ile Cys Thr Ile Gln Ser Ser Gly Gly Ile Met Leu Leu Lys
145     150     155     160
Asp Gln Val Leu Arg Cys Ser Lys Ile Ala Gly Val Lys Val Ala Glu
165     170     175
Ile Thr Glu Leu Ile Leu Lys Ala Leu Glu Asn Asp Gln Lys Val Arg
180     185     190
Lys Glu Gly Gly Lys Phe Gly Phe Ala Glu Ser Ile Ala Asn Gln Arg
195     200     205
Ile Thr Ala Phe Lys Met Glu Lys Ala Pro Ile Asp Thr Ser Asp Val
210     215     220
Glu Glu Lys Ala Glu Glu Ile Ile Ala Glu Ala Glu Pro Pro Ser Glu
225     230     235     240
Val Val Ser Thr Pro Val Leu Trp Thr Pro Gly Thr Ala Gln Ile Gly
245     250     255
Glu Gly Val Glu Asn Ser Trp Gly Asp Leu Glu Asp Ser Glu Lys Glu
260     265     270
Asp Asp Glu Gly Gly Gly Asp Gln Ala Ile Ile Leu Asp Gly Ile Lys
275     280     285
Met Asp Thr Gly Val Glu Val Ser Asp Ile Gly Ser Gln Asp Ala Pro
290     295     300
Ile Ile Leu Ser Asp Ser Glu Glu Glu Glu Met Ile Ile Leu Glu Pro
305     310     315     320
Asp Lys Asn Pro Lys Lys Ile Arg Thr Gln Thr Thr Ser Ala Lys Gln
325     330     335
Glu Lys Ala Pro Ser Lys Lys Pro Val Lys Arg Arg Lys Lys Lys Arg
340     345     350
Ala Ala Asn
355

```

5

<210> 11

<211> 664
 <212> PRT
 <213> *Homo sapiens*

5 <400> 11

```

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

Ile Pro Thr Asn Ala Arg Leu Leu Ser Asp His Ser Lys Pro Thr Ala
20           25           30

Glu Thr Val Ala Pro Asp Asn Thr Ala Ile Pro Ser Leu Trp Ala Glu
35           40           45

Ala Glu Glu Asn Glu Lys Glu Thr Ala Val Ser Thr Glu Asp Asp Ser
50           55           60

His His Lys Ala Glu Lys Ser Ser Val Leu Lys Ser Lys Glu Glu Ser
65           70           75           80

His Glu Gln Ser Ala Glu Gln Gly Lys Ser Ser Ser Gln Glu Leu Gly
85           90           95

Leu Lys Asp Gln Glu Asp Ser Asp Gly His Leu Ser Val Asn Leu Glu
100          105          110

Tyr Ala Pro Thr Glu Gly Thr Leu Asp Ile Lys Glu Asp Met Ile Glu
115          120          125

Pro Gln Glu Lys Lys Leu Ser Glu Asn Thr Asp Phe Leu Ala Pro Gly
130          135          140

Val Ser Ser Phe Thr Asp Ser Asn Gln Gln Glu Ser Ile Thr Lys Arg
145          150          155          160

Glu Glu Asn Gln Glu Gln Pro Arg Asn Tyr Ser His His Gln Leu Asn
165          170          175

Arg Ser Ser Lys His Ser Gln Gly Leu Arg Asp Gln Gly Asn Gln Glu
180          185          190

```

Gln Asp Pro Asn Ile Ser Asn Gly Glu Glu Glu Glu Glu Lys Glu Pro
 195 200 205
 Gly Glu Val Gly Thr His Asn Asp Asn Gln Glu Arg Lys Thr Glu Leu
 210 215 220
 Pro Arg Glu His Ala Asn Ser Lys Gln Glu Glu Asp Asn Thr Gln Ser
 225 230 235 240
 Asp Asp Ile Leu Glu Glu Ser Asp Gln Pro Thr Gln Val Ser Lys Met
 245 250 255
 Gln Glu Asp Glu Phe Asp Gln Gly Asn Gln Glu Gln Glu Asp Asn Ser
 260 265 270
 Asn Ala Glu Met Glu Glu Glu Asn Ala Ser Asn Val Asn Lys His Ile
 275 280 285
 Gln Glu Thr Glu Trp Gln Ser Gln Glu Gly Lys Thr Gly Leu Glu Ala
 290 295 300
 Ile Ser Asn His Lys Glu Thr Glu Glu Lys Thr Val Ser Glu Ala Leu
 305 310 315 320
 Leu Met Glu Pro Thr Asp Asp Gly Asn Thr Thr Pro Arg Asn His Gly
 325 330 335
 Val Asp Asp Asp Gly Asp Asp Asp Gly Asp Asp Gly Gly Thr Asp Gly
 340 345 350
 Pro Arg His Ser Ala Ser Asp Asp Tyr Phe Ile Pro Ser Gln Ala Phe
 355 360 365
 Leu Glu Ala Glu Arg Ala Gln Ser Ile Ala Tyr His Leu Lys Ile Glu
 370 375 380
 Glu Gln Arg Glu Lys Val His Glu Asn Glu Asn Ile Gly Thr Thr Glu
 385 390 395 400
 Pro Gly Glu His Gln Glu Ala Lys Lys Ala Glu Asn Ser Ser Asn Glu
 405 410 415
 Glu Glu Thr Ser Ser Glu Gly Asn Met Arg Val His Ala Val Asp Ser
 420 425 430
 Cys Met Ser Phe Gln Cys Lys Arg Gly His Ile Cys Lys Ala Asp Gln
 435 440 445
 Gln Gly Lys Pro His Cys Val Cys Gln Asp Pro Val Thr Cys Pro Pro
 450 455 460
 Thr Lys Pro Leu Asp Gln Val Cys Gly Thr Asp Asn Gln Thr Tyr Ala
 465 470 475 480
 Ser Ser Cys His Leu Phe Ala Thr Lys Cys Arg Leu Glu Gly Thr Lys
 485 490 495
 Lys Gly His Gln Leu Gln Leu Asp Tyr Phe Gly Ala Cys Lys Ser Ile
 500 505 510
 Pro Thr Cys Thr Asp Phe Glu Val Ile Gln Phe Pro Leu Arg Met Arg
 515 520 525

Asp Trp Leu Lys Asn Ile Leu Met Gln Leu Tyr Glu Ala Asn Ser Glu
 530 535 540
 His Ala Gly Tyr Leu Asn Glu Lys Gln Arg Asn Lys Val Lys Lys Ile
 545 550 555 560
 Tyr Leu Asp Glu Lys Arg Leu Leu Ala Gly Asp His Pro Ile Asp Leu
 565 570 575
 Leu Leu Arg Asp Phe Lys Lys Asn Tyr His Met Tyr Val Tyr Pro Val
 580 585 590
 His Trp Gln Phe Ser Glu Leu Asp Gln His Pro Met Asp Arg Val Leu
 595 600 605
 Thr His Ser Glu Leu Ala Pro Leu Arg Ala Ser Leu Val Pro Met Glu
 610 615 620
 His Cys Ile Thr Arg Phe Phe Glu Glu Cys Asp Pro Asn Lys Asp Lys
 625 630 635 640
 His Ile Thr Leu Lys Glu Trp Gly His Cys Phe Gly Ile Lys Glu Glu
 645 650 655
 Asp Ile Asp Glu Asn Leu Leu Phe
 660

<210> 12
 <211> 664
 <212> PRT
 <213> *Homo sapiens*

<400> 12

Met Lys Thr Gly Leu Phe Phe Leu Cys Leu Leu Gly Thr Ala Ala Ala
 1 5 10 15
 Ile Pro Thr Asn Ala Arg Leu Leu Ser Asp His Ser Lys Pro Thr Ala
 20 25 30
 Glu Thr Val Ala Pro Asp Asn Thr Ala Ile Pro Ser Leu Arg Ala Glu
 35 40 45
 Asp Glu Glu Asn Glu Lys Glu Thr Ala Val Ser Thr Glu Asp Asp Ser
 50 55 60
 His His Lys Ala Glu Lys Ser Ser Val Leu Lys Ser Lys Glu Glu Ser
 65 70 75 80
 His Glu Gln Ser Ala Glu Gln Gly Lys Ser Ser Ser Gln Glu Leu Gly
 85 90 95
 Leu Lys Asp Gln Glu Asp Ser Asp Gly Asp Leu Ser Val Asn Leu Glu
 100 105 110
 Tyr Ala Pro Ser Glu Gly Thr Leu Asp Ile Lys Glu Asp Met Ser Glu
 115 120 125
 Pro Gln Glu Lys Lys Leu Ser Glu Asn Thr Asp Phe Leu Ala Pro Gly
 130 135 140

Val	Ser	Ser	Phe	Thr	Asp	Ser	Asn	Gln	Gln	Glu	Ser	Ile	Thr	Lys	Arg		
145					150					155					160		
Glu	Glu	Asn	Gln	Glu	Gln	Pro	Arg	Asn	Tyr	Ser	His	His	Gln	Leu	Asn		
				165					170					175			
Arg	Ser	Ser	Lys	His	Ser	Gln	Gly	Leu	Arg	Asp	Gln	Gly	Asn	Gln	Glu		
			180					185					190				
Gln	Asp	Pro	Asn	Ile	Ser	Asn	Gly	Glu	Glu	Glu	Glu	Glu	Lys	Glu	Pro		
		195					200						205				
Gly	Glu	Val	Gly	Thr	His	Asn	Asp	Asn	Gln	Glu	Arg	Lys	Thr	Glu	Leu		
210						215					220						
Pro	Arg	Glu	His	Ala	Asn	Ser	Lys	Gln	Glu	Glu	Asp	Asn	Thr	Gln	Ser		
225					230					235					240		
Asp	Asp	Ile	Leu	Glu	Glu	Ser	Asp	Gln	Pro	Thr	Gln	Val	Ser	Lys	Met		
				245					250					255			
Gln	Glu	Asp	Glu	Phe	Asp	Gln	Gly	Asn	Gln	Glu	Gln	Glu	Asp	Asn	Ser		
			260					265					270				
Asn	Ala	Glu	Met	Glu	Glu	Glu	Asn	Ala	Ser	Asn	Val	Asn	Lys	His	Ile		
		275					280					285					
Gln	Glu	Thr	Glu	Trp	Gln	Ser	Gln	Glu	Gly	Lys	Thr	Gly	Leu	Glu	Ala		
290					295						300						
Ile	Ser	Asn	His	Lys	Glu	Thr	Glu	Glu	Lys	Thr	Val	Ser	Glu	Ala	Leu		
305				310						315					320		
Leu	Met	Glu	Pro	Thr	Asp	Asp	Gly	Asn	Thr	Thr	Pro	Arg	Asn	His	Gly		
				325					330					335			
Val	Asp	Asp	Asp	Gly	Asp	Asp	Asp	Gly	Asp	Asp	Gly	Gly	Thr	Asp	Gly		
			340					345					350				
Pro	Arg	His	Ser	Ala	Ser	Asp	Asp	Tyr	Phe	Ile	Pro	Ser	Gln	Ala	Phe		
		355					360					365					
Leu	Glu	Ala	Glu	Arg	Ala	Gln	Ser	Ile	Ala	Tyr	His	Leu	Lys	Ile	Glu		
370						375					380						
Glu	Gln	Arg	Glu	Lys	Val	His	Glu	Asn	Glu	Asn	Ile	Gly	Thr	Thr	Glu		
385				390						395					400		
Pro	Gly	Glu	His	Gln	Glu	Ala	Lys	Lys	Ala	Glu	Asn	Ser	Ser	Asn	Glu		
				405					410					415			
Glu	Glu	Thr	Ser	Ser	Glu	Gly	Asn	Met	Arg	Val	His	Ala	Val	Asp	Ser		
			420					425					430				
Cys	Met	Ser	Phe	Gln	Cys	Lys	Arg	Gly	His	Ile	Cys	Lys	Ala	Asp	Gln		
		435					440					445					
Gln	Gly	Lys	Pro	His	Cys	Val	Cys	Gln	Asp	Pro	Val	Thr	Cys	Pro	Pro		
450						455					460						
Thr	Lys	Pro	Leu	Asp	Gln	Val	Cys	Gly	Thr	Asp	Asn	Gln	Thr	Tyr	Ala		
465					470					475					480		

Ser Ser Cys His Leu Phe Ala Thr Lys Cys Arg Leu Glu Gly Thr Lys
485 490 495

Lys Gly His Gln Leu Gln Leu Asp Tyr Phe Gly Ala Cys Lys Ser Ile
500 505 510

Pro Thr Cys Thr Asp Phe Glu Val Ile Gln Phe Pro Leu Arg Met Arg
515 520 525

Asp Trp Leu Lys Asn Ile Leu Met Gln Leu Tyr Glu Ala Asn Ser Glu
530 535 540

His Ala Gly Tyr Leu Asn Glu Lys Gln Arg Asn Lys Val Lys Lys Ile
545 550 555 560

Tyr Leu Asp Glu Lys Arg Leu Leu Ala Gly Asp His Pro Ile Asp Leu
565 570 575

Leu Leu Arg Asp Phe Lys Lys Asn Tyr His Met Tyr Val Tyr Pro Val
580 585 590

His Trp Gln Phe Ser Glu Leu Asp Gln His Pro Met Asp Arg Val Leu
595 600 605

Thr His Ser Glu Leu Ala Pro Leu Arg Ala Ser Leu Val Pro Met Glu
610 615 620

His Cys Ile Thr Arg Phe Phe Glu Glu Cys Asp Pro Asn Lys Asp Lys
625 630 635 640

His Ile Thr Leu Lys Glu Trp Gly His Cys Phe Gly Ile Lys Glu Glu
645 650 655

Asp Ile Asp Glu Asn Leu Leu Phe
660

<210> 13
<211> 664
<212> PRT
<213> *Homo sapiens*

5

<400> 13

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

Ile Pro Thr Asn Ala Arg Leu Leu Ser Asp His Ser Lys Pro Thr Ala
20 25 30

Glu Thr Val Ala Pro Asp Asn Thr Ala Ile Pro Ser Leu Trp Ala Glu
35 40 45

Ala Glu Glu Asn Glu Lys Glu Thr Ala Val Ser Thr Glu Asp Asp Ser
50 55 60

His His Lys Ala Glu Lys Ser Ser Val Leu Lys Ser Lys Glu Glu Ser
65 70 75 80

His Glu Gln Ser Ala Glu Gln Gly Lys Ser Ser Ser Gln Glu Leu Gly
85 90 95

Leu Lys Asp Gln Glu Asp Ser Asp Gly His Leu Ser Val Asn Leu Glu
 100 105 110
 Tyr Ala Pro Thr Glu Gly Thr Leu Asp Ile Lys Glu Asp Met Ile Glu
 115 120 125
 Pro Gln Glu Lys Lys Leu Ser Glu Asn Thr Asp Phe Leu Ala Pro Gly
 130 135 140
 Val Ser Ser Phe Thr Asp Ser Asn Gln Gln Glu Ser Ile Thr Lys Arg
 145 150 155 160
 Glu Glu Asn Gln Glu Gln Pro Arg Asn Tyr Ser His His Gln Leu Asn
 165 170 175
 Arg Ser Ser Lys His Ser Gln Gly Leu Arg Asp Gln Gly Asn Gln Glu
 180 185 190
 Gln Asp Pro Asn Ile Ser Asn Gly Glu Glu Glu Glu Glu Lys Glu Pro
 195 200 205
 Gly Glu Val Gly Thr His Asn Asp Asn Gln Glu Arg Lys Thr Glu Leu
 210 215 220
 Pro Arg Glu His Ala Asn Ser Lys Gln Glu Glu Asp Asn Thr Gln Ser
 225 230 235 240
 Asp Asp Ile Leu Glu Glu Ser Asp Gln Pro Thr Gln Val Ser Lys Met
 245 250 255
 Gln Glu Asp Glu Phe Asp Gln Gly Asn Gln Glu Gln Glu Asp Asn Ser
 260 265 270
 Asn Ala Glu Met Glu Glu Glu Asn Ala Ser Asn Val Asn Lys His Ile
 275 280 285
 Gln Glu Thr Glu Trp Gln Ser Gln Glu Gly Lys Thr Gly Leu Glu Ala
 290 295 300
 Ile Ser Asn His Lys Glu Thr Glu Glu Lys Thr Val Ser Glu Ala Leu
 305 310 315 320
 Leu Met Glu Pro Thr Asp Asp Gly Asn Thr Thr Pro Arg Asn His Gly
 325 330 335
 Val Asp Asp Asp Gly Asp Asp Asp Gly Asp Asp Gly Gly Thr Asp Gly
 340 345 350
 Pro Arg His Ser Ala Ser Asp Asp Tyr Phe Ile Pro Ser Gln Ala Phe
 355 360 365
 Leu Glu Ala Glu Arg Ala Gln Ser Ile Ala Tyr His Leu Lys Ile Glu
 370 375 380
 Glu Gln Arg Glu Lys Val His Glu Asn Glu Asn Ile Gly Thr Thr Glu
 385 390 395 400
 Pro Gly Glu His Gln Glu Ala Lys Lys Ala Glu Asn Ser Ser Asn Glu
 405 410 415
 Glu Glu Thr Ser Ser Glu Gly Asn Met Arg Val His Ala Val Asp Ser
 420 425 430

Cys Met Ser Phe Gln Cys Lys Arg Gly His Ile Cys Lys Ala Asp Gln
435 440 445

Gln Gly Lys Pro His Cys Val Cys Gln Asp Pro Val Thr Cys Pro Pro
450 455 460

Thr Lys Pro Leu Asp Gln Val Cys Gly Thr Asp Asn Gln Thr Tyr Ala
465 470 475 480

Ser Ser Cys His Leu Phe Ala Thr Lys Cys Arg Leu Glu Gly Thr Lys
485 490 495

Lys Gly His Gln Leu Gln Leu Asp Tyr Phe Gly Ala Cys Lys Ser Ile
500 505 510

Pro Thr Cys Thr Asp Phe Glu Val Ile Gln Phe Pro Leu Arg Met Arg
515 520 525

Asp Trp Leu Lys Asn Ile Leu Met Gln Leu Tyr Glu Ala Asn Ser Glu
530 535 540

His Ala Gly Tyr Leu Asn Glu Lys Gln Arg Asn Lys Val Lys Lys Ile
545 550 555 560

Tyr Leu Asp Glu Lys Arg Leu Leu Ala Gly Asp His Pro Ile Asp Leu
565 570 575

Leu Leu Arg Asp Phe Lys Lys Asn Tyr His Met Tyr Val Tyr Pro Val
580 585 590

His Trp Gln Phe Ser Glu Leu Asp Gln His Pro Met Asp Arg Val Leu
595 600 605

Thr His Ser Glu Leu Ala Pro Leu Arg Ala Ser Leu Val Pro Met Glu
610 615 620

His Cys Ile Thr Arg Phe Phe Glu Glu Cys Asp Pro Asn Lys Asp Lys
625 630 635 640

His Ile Thr Leu Lys Glu Trp Gly His Cys Phe Gly Ile Lys Glu Glu
645 650 655

Asp Ile Asp Glu Asn Leu Leu Phe
660

<210> 14
<211> 676
<212> PRT
<213> *Coturnix coturnix*

<400> 14

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

Thr Pro Thr Asp Pro Leu Asn Tyr Gln Phe Gly Ala His Gly Gln Lys
20 25 30

Thr Ala Glu Lys His Lys Tyr Thr His Ser Glu Met Pro Glu Glu Glu
35 40 45

Asn Thr Gly Phe Val Asn Lys Gly Asp Val Leu Ser Gly His Arg Thr
 50 55 60
 Ile Lys Ala Glu Val Pro Val Leu Asp Thr Gln Lys Asp Glu Pro Trp
 65 70 75 80
 Ala Ser Arg Arg Gln Gly Gln Gly Asp Gly Glu His Gln Thr Lys Asn
 85 90 95
 Ser Leu Arg Ser Ile Asn Phe Leu Thr Leu His Ser Asn Pro Gly Leu
 100 105 110
 Ala Ser Asp Asn Gln Glu Ser Asn Ser Gly Ser Ser Arg Glu Gln His
 115 120 125
 Ser Ser Glu His His Gln Pro Arg Arg His Arg Lys His Gly Asn Met
 130 135 140
 Ala Gly Gln Trp Ala Leu Arg Gly Glu Ser Pro Val Asp Ala Leu Gly
 145 150 155 160
 Leu Val Arg Glu Arg Asn Thr Trp Lys Tyr Asn Lys Asn Thr Val Gly
 165 170 175
 Leu Asp Glu Asn Asn Asn Gly Ser Glu Glu Glu Glu Ala Gly Glu Glu
 180 185 190
 Glu Asp Glu Glu Trp Gly Glu Glu Thr Asp Tyr Arg Asp Met Lys His
 195 200 205
 Arg Ala Arg Gly Thr Ser His Gly Arg Glu Tyr Arg Arg Trp Gln Asn
 210 215 220
 Glu Asn Ser Arg Pro Ser Gly Glu Phe Leu Arg Asp Ser Ser Leu Pro
 225 230 235 240
 Val Arg Ile Thr Lys Arg His Gly Glu Lys Phe Ser Met Glu Glu Glu
 245 250 255
 Ser Gln Glu Lys Leu Tyr Lys Glu Gly Lys Leu Pro Leu Ser Lys Lys
 260 265 270
 Asn His Asn Glu Asp Gln Gly Glu Lys Arg Gln Ser Glu Glu Ser Lys
 275 280 285
 Glu His Phe Gln Val Val Asn Gln Arg Lys His Arg Ala Val Thr Lys
 290 295 300
 Arg Gln Asp Lys Glu Gly Ser Asn Ala Glu Glu Asp Asp Asn Asp Ser
 305 310 315 320
 Gly Asp Asp Gly Glu Glu Asp Leu Gly Asn Val Trp Arg Glu Ala Val
 325 330 335
 Tyr Glu Glu Glu Glu Arg Met Gln Ser Asn Asp Gln Asp Ser Ile Thr
 340 345 350
 Asn Lys Gln Lys Glu Glu Ile Thr Ala Gly Asp Asp Ser Gly Val Tyr
 355 360 365
 Arg Glu Met Gln Asp Tyr Lys Gly Asp Lys Ile Lys Asp Val Thr His
 370 375 380

Ser Glu Asp Asn His Tyr His His Glu Pro Pro Asn Ser Ser Ser Lys
385 390 395 400

Gln Gln Leu Gln Thr Ser Ser Ser Val Glu Ser Met Asn Ser Thr Glu
405 410 415

His Glu Asp Glu Val Lys Thr Thr Gly Gly Ser Tyr His Glu Glu Ser
420 425 430

Ala Arg Asn Ser Thr Gly Lys Ala Leu Pro Asp Leu Cys Arg Asn Phe
435 440 445

His Cys Lys Arg Gly Lys Val Cys Gln Ala Asp Lys Gln Gly Lys Pro
450 455 460

Ser Cys Ile Cys Gln Asp Pro Ala Ala Cys Pro Ser Thr Lys Asp Tyr
465 470 475 480

Lys Arg Val Cys Gly Thr Asp Asn Lys Thr Tyr Asp Gly Thr Cys Gln
485 490 495

Leu Phe Gly Thr Lys Cys Gln Leu Glu Gly Thr Lys Met Gly Arg Gln
500 505 510

Leu His Leu Asp Tyr Met Gly Ala Cys Lys His Ile Pro His Cys Thr
515 520 525

Asp Tyr Glu Val Asn Gln Phe Pro Leu Arg Met Arg Asp Trp Leu Lys
530 535 540

Asn Ile Leu Met Gln Tyr Tyr Glu Arg Asp Gln Asp Thr Ser Ala Phe
545 550 555 560

Leu Thr Glu Lys Gln Arg Asn Lys Val Lys Lys Ile Tyr Leu Asn Glu
565 570 575

Lys Arg Leu Val Ser Gly Glu His Pro Val Glu Leu Leu Leu His Asp
580 585 590

Phe Glu Lys Asn Tyr His Met Tyr Leu Tyr Pro Val His Trp Gln Phe
595 600 605

Tyr Gln Leu Asp Gln His Pro Val Asp Arg Ser Leu Thr His Ser Glu
610 615 620

Leu Ala Pro Leu Arg Ala Ser Leu Val Pro Met Glu His Cys Ile Thr
625 630 635 640

Arg Phe Phe Gln Glu Cys Asp Gly Asp Gln Asp Lys Leu Ile Thr Leu
645 650 655

Lys Glu Trp Cys His Cys Phe Ala Ile Lys Glu Glu Asp Ile Asn Glu
660 665 670

Asn Leu Leu Phe
675

<210> 15
<211> 306
<212> PRT
<213> *Mus musculus*
<400> 15

Met Trp Lys Arg Trp Leu Ala Leu Ser Leu Val Thr Ile Ala Leu Val
1 5 10 15
His Gly Glu Glu Glu Pro Arg Ser Lys Ser Lys Ile Cys Ala Asn Val
20 25 30
Phe Cys Gly Ala Gly Arg Glu Cys Ala Val Thr Glu Lys Gly Glu Pro
35 40 45
Thr Cys Leu Cys Ile Glu Gln Cys Lys Pro His Lys Arg Pro Val Cys
50 55 60
Gly Ser Asn Gly Lys Thr Tyr Leu Asn His Cys Glu Leu His Arg Asp
65 70 75 80
Ala Cys Leu Thr Gly Ser Lys Ile Gln Val Asp Tyr Asp Gly His Cys
85 90 95
Lys Glu Lys Lys Ser Ala Ser Pro Ser Ala Ser Pro Val Val Cys Tyr
100 105 110
Gln Ala Asn Arg Asp Glu Leu Arg Arg Arg Leu Ile Gln Trp Leu Glu
115 120 125
Ala Glu Ile Ile Pro Asp Gly Trp Phe Ser Lys Gly Ser Asn Tyr Ser
130 135 140
Glu Ile Leu Asp Lys Tyr Phe Lys Ser Phe Asp Asn Gly Asp Ser His
145 150 155 160
Leu Asp Ser Ser Glu Phe Leu Lys Phe Val Glu Gln Asn Glu Thr Ala
165 170 175
Ile Asn Ile Thr Thr Tyr Ala Asp Gln Glu Asn Asn Lys Leu Leu Arg
180 185 190
Ser Leu Cys Val Asp Ala Leu Ile Glu Leu Ser Asp Glu Asn Ala Asp
195 200 205
Trp Lys Leu Ser Phe Gln Glu Phe Leu Lys Cys Leu Asn Pro Ser Phe
210 215 220
Asn Pro Pro Glu Lys Lys Cys Ala Leu Glu Asp Glu Thr Tyr Ala Asp
225 230 235 240
Gly Ala Glu Thr Glu Val Asp Cys Asn Arg Cys Val Cys Ser Cys Gly
245 250 255
His Trp Val Cys Thr Ala Met Thr Cys Asp Gly Lys Asn Gln Lys Gly
260 265 270
Val Gln Thr His Thr Glu Glu Glu Lys Thr Gly Tyr Val Gln Glu Leu
275 280 285
Gln Lys His Gln Gly Thr Ala Glu Lys Thr Lys Lys Val Asn Thr Lys
290 295 300
Glu Ile
305

<210> 16
<211> 439
<212> PRT
<213> *Homo sapiens*

<400> 16

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

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

Pro Asn His Gly Asn Phe Leu Asp Asn Asp Gln Trp Leu Ser Thr Val
35 40 45

Ser Gln Tyr Asp Arg Asp Lys Tyr Trp Asn Arg Phe Arg Asp Asp Asp
50 55 60

Tyr Phe Arg Asn Trp Asn Pro Asn Lys Pro Phe Asp Gln Ala Leu Asp
65 70 75 80

Pro Ser Lys Asp Pro Cys Leu Lys Val Lys Cys Ser Pro His Lys Val
85 90 95

Cys Val Thr Gln Asp Tyr Gln Thr Ala Leu Cys Val Ser Arg Lys His
100 105 110

Leu Leu Pro Arg Gln Lys Lys Gly Asn Val Ala Gln Lys His Trp Val
115 120 125

Gly Pro Ser Asn Leu Val Lys Cys Lys Pro Cys Pro Val Ala Gln Ser
130 135 140

Ala Met Val Cys Gly Ser Asp Gly His Ser Tyr Thr Ser Lys Cys Lys
145 150 155 160

Leu Glu Phe His Ala Cys Ser Thr Gly Lys Ser Leu Ala Thr Leu Cys
165 170 175

Asp Gly Pro Cys Pro Cys Leu Pro Glu Pro Glu Pro Pro Lys His Lys
180 185 190

Ala Glu Arg Ser Ala Cys Thr Asp Lys Glu Leu Arg Asn Leu Ala Ser
195 200 205

Arg Leu Lys Asp Trp Phe Gly Ala Leu His Glu Asp Ala Asn Arg Val
210 215 220

Ile Lys Pro Thr Ser Ser Asn Thr Ala Gln Gly Arg Phe Asp Thr Ser
225 230 235 240

Ile Leu Pro Ile Cys Lys Asp Ser Leu Gly Trp Met Phe Asn Lys Leu
245 250 255

Asp Met Asn Tyr Asp Leu Leu Leu Asp Pro Ser Glu Ile Asn Ala Ile
260 265 270

Tyr Leu Asp Lys Tyr Glu Pro Cys Ile Lys Pro Leu Phe Asn Ser Cys
275 280 285

Asp Ser Phe Lys Asp Gly Lys Leu Ser Asn Asn Glu Trp Cys Tyr Cys

290 295 300
 Phe Gln Lys Pro Gly Gly Leu Pro Cys Gln Asn Glu Met Asn Arg Ile
 305 310 315 320
 Gln Lys Leu Ser Lys Gly Lys Ser Leu Leu Gly Ala Phe Ile Pro Arg
 325 330 335
 Cys Asn Glu Glu Gly Tyr Tyr Lys Ala Thr Gln Cys His Gly Ser Thr
 340 345 350
 Gly Gln Cys Trp Cys Val Asp Lys Tyr Gly Asn Glu Leu Ala Gly Ser
 355 360 365
 Arg Lys Gln Gly Ala Val Ser Cys Glu Glu Glu Gln Glu Thr Ser Gly
 370 375 380
 Asp Phe Gly Ser Gly Gly Ser Val Val Leu Leu Asp Asp Leu Glu Tyr
 385 390 395 400
 Glu Arg Glu Leu Gly Pro Lys Asp Lys Glu Gly Lys Leu Arg Val His
 405 410 415
 Thr Arg Ala Val Thr Glu Asp Asp Glu Asp Glu Asp Asp Asp Lys Glu
 420 425 430
 Asp Glu Val Gly Tyr Ile Trp
 435

<210> 17
 <211> 263
 <212> PRT
 <213> *Mus musculus*

<400> 17

Met Leu Pro Cys Phe Gln Leu Leu Arg Ile Gly Gly Gly Arg Gly Gly
 1 5 10 15
 Asp Leu Tyr Thr Phe His Pro Pro Ser Lys Ser Gly Cys Thr Tyr Gly
 20 25 30
 Leu Gly Cys Arg Ala Asp Leu Cys Asp Val Ala Leu Arg Pro Gln Gln
 35 40 45
 Glu Pro Gly Leu Ile Ser Gly Val His Ala Glu Leu His Ala Glu Leu
 50 55 60
 Gln Gly Asp Asp Trp Arg Val Ser Leu Glu Asp His Ser Ser Gln Gly
 65 70 75 80
 Thr Leu Val Asn Asn Val Arg Leu Pro Arg Gly His Arg Leu Glu Leu
 85 90 95
 Ser Asp Gly Asp Leu Leu Thr Phe Gly Pro Gln Gly Gln Ala Gly Thr
 100 105 110
 Ser Ser Ser Ser Glu Phe Tyr Phe Met Phe Gln Gln Val Arg Val Lys
 115 120 125
 Pro Gln Asp Phe Ala Ala Ile Thr Val Pro Arg Ser Lys Gly Glu Ala
 130 135 140

Gly Gly Gly Phe Gln Pro Met Leu Pro Pro Gln Gly Ala Pro Gln Arg
 145 150 155 160
 Pro Leu Ser Thr Leu Ser Ser Ala Pro Lys Ala Thr Leu Ile Leu Asn
 165 170 175
 Ser Ile Gly Ser Leu Ser Lys Leu Gln Ala Gln Pro Leu Thr Phe Ser
 180 185 190
 Arg Gly Gly Gly Arg Pro Gln Gly Leu Ala Ile Pro Ser Gln His Gly
 195 200 205
 Glu Ala Gln Val Ser Pro Ala Pro Pro Thr Arg Asn Arg Arg Lys Ser
 210 215 220
 Ala His Lys Val Leu Ala Glu Leu Asp Asp Glu Val Ser Pro Gly Pro
 225 230 235 240
 Leu Ser Val Leu Thr Glu Pro Arg Lys Arg Leu Arg Val Glu Lys Ala
 245 250 255
 Ala Leu Ile Ala Ser Gly Glu
 260

<210> 18
 <211> 2133
 <212> ADN
 <213> *Homo sapiens*

<400> 18

cgggagagcg cgctctgcct gccgcctgcc tgccctgccac tgagggttcc cagcaccatg 60
 agggcctgga tcttctttct cctttgcctg gccgggaggg ccttggcagc ccctcagcaa 120
 gaagccctgc ctgatgagac agaggtggtg gaagaaactg tggcagaggt gactgaggtg 180
 tctgtgggag ctaatcctgt ccaggtggaa gtaggagaat ttgatgatgg tgcagaggaa 240
 accgaagagg aggtggtggc ggaaaatccc tgccagaacc accactgcaa acacggcaag 300
 gtgtgcgagc tggatgagaa caacaccccc atgtgcgtgt gccaggaccc caccagctgc 360
 ccagccccca ttggcgagtt tgagaagggtg tgcagcaatg acaacaagac cttcgactct 420
 tcctgccact tctttgccac aaagtgcacc ctggagggca ccaagaaggg ccacaagctc 480
 cacctggact acatcgggcc ttgcaaatac atccccctt gcctggactc tgagctgacc 540
 gaattcccc tgccatgcg ggactggctc aagaacgtcc tggtcaccct gtatgagagg 600
 gatgaggaca acaaccttct gactgagaag cagaagctgc gggatgaagaa gatccatgag 660
 aatgagaagc gcctggaggc aggagaccac cccgtggagc tgctggcccg ggacttegag 720
 aagaactata acatgtacat ctccctgta cactggcagt tcggccagct ggaccagcac 780
 cccattgacg ggtacctctc ccacaccgag ctggctccac tgcgtgctcc cctcatcccc 840

ES 2 565 496 T3

```

atggagcatt gcaccacccg ctttttcgag acctgtgacc tggacaatga caagtacatc 900
gccctggatg agtgggcccg ctgcttcggc atcaagcaga aggatatcga caaggatctt 960
gtgatctaaa tccactcctt ccacagtacc ggattctctc tttaacccctc cccttcgtgt 1020
ttcccccaat gtttaaaatg tttggatggt ttgttgttct gcctggagac aagggtgctaa 1080
catagattta agtgaatata ttaacgggtc taaaaatgaa aattctaacc caagacatga 1140
cattcttagc tgtaacttaa ctattaaggc cttttccaca cgcattaata gtcccatttt 1200
tctcttgcca tttgtagctt tgcccattgt cttattggca catgggtgga cacggatctg 1260
ctgggctctg ccttaaacac acattgcagc ttcaactttt ctcttttagt ttctgtttga 1320
aactaatact taccgagtca gactttgtgt tcatttcatt tcagggtctt ggctgcctgt 1380
gggcttcccc aggtggcctg gaggtgggca aaggaagta acagacacac gatgttgca 1440
aggatggttt tgggactaga ggctcagtg tgggagagat ccctgcagaa tccaccaacc 1500
agaacgtggt ttgcctgagg ctgtaactga gagaagatt ctggggctgt cttatgaaaa 1560
tatagacatt ctcacataag cccagttcat caccatttcc tcctttacct ttcagtgcag 1620
tttcttttca cattaggctg ttggttcaaa cttttgggag cacggactgt cagttctctg 1680
ggaagtggtc agcgcatact gcagggttc tctcctctg tcttttgag aaccagggt 1740
cttctcaggg gctctagga ctgccaggct gtttcagcca ggaaggcca aatcaagat 1800
gagatgtaga aagttgtaaa atagaaaaag tggagtggg gaatcggtt ttctttcctc 1860
acatttggat gattgtcata aggtttttag catgttcctc cttttcttca ccctcccctt 1920
tgttctteta ttaatcaaga gaaacttcaa agttaatgg atggtcggat ctcacaggct 1980
gagaactcgt tcacctcaa gcatttcag aaaaagctgc ttcttattaa tcatacaaac 2040
tctcaccatg atgtgaagag tttcacaaat ctttcaaaat aaaaagtaat gacttagaaa 2100
ctgaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 2133

```

<210> 19
 <211> 2079
 <212> ADN
 <213> *Mus musculus*

<400> 19

```

gcattcctgc agcccttcag accgccagaa ctcttctgcc gcctgcctgc ctgcctgcct 60
gcctgcctgt gccgagagt cccagcatca tgagggcctg gatcttcttt ctcccttgcc 120
tgcccgaggag ggccctggca gcccctcagc agactgaagt tgctgaggag atagtggagg 180
aggaaaccgt ggtggaggag acagggttac ctgtgggtgc caaccagtc caggtgga 240
tgggagaatt tgaggacggt gcagaggaaa cggctgagga ggtggtggct gacaaccct 300

```

```

gccagaacca tcattgcaaa catggcaagg tgtgtgagct ggacgagagc aacacccccca 360
tgtgtgtgtg ccaggacccc accagctgcc ctgctcccat tggcgagttt gagaaggtat 420
gcagcaatga caacaagacc ttcgactcct cctgccactt ctttgccacc aagtgcaccc 480
tggagggcac caagaagggc cacaagctcc acctggacta catcggacca tgcaaatata 540
tcgccccctg cctggattcc gagctgaccg aattccctct gcgcatgcgt gactggctca 600
aaaatgtcct ggtcaccttg tacgagagag atgagggcaa caacctctc actgagaagc 660
agaagctgcg tgtgaagaag atccatgaga atgagaagcg cctggaggct ggagaccacc 720
ccgtggagct gttggcccga gactttgaga agaactacaa tatgtacatc ttccctgtcc 780
actggcagtt tggccagctg gatcagcacc ctattgatgg gtacctgtcc cacttgagc 840
tgccccact gcgtgctccc ctcatcccca tggaacattg caccacacgt ttctttgaga 900
cctgtgacct agacaacgac aagtacattg ccctggagga atgggcccgc tgctttggca 960
tcaaggagca ggacatcaac aaggatctgg tgatctaagt tcacgcctcc tgcgtcagtc 1020
ctgaactctc tccctctgat gtgtcacccc tccattacc ccctgttta aaatgtttgg 1080
atggttggct gttccgcctg gggataaggt gctaacatag atttaactga atacattaac 1140
ggtgctaaaa aaaaaaaaaa aacaaagtaa gaaagaaact agaaccaag tcacagcatt 1200
ttcccacata actctgaggc catggcccat ccacagcctc ctggccccct gcactacca 1260
gtgtctcact ggctgtgttg gaaacggagt tgcataagct caccgtccac aagcacgaga 1320
tatctctagc tttcatttca attttgcatt tgactcttaa cactcaccca gactctgtgc 1380
ttatttcatt ttgggggatg tgggcttttt cccctggtgg tttggagtta ggcagaggga 1440
agttacagac acaggtacaa aatttgggta aagatactgt gagacctgag gaccaccag 1500
tcagaaccca catggcaagt cttagtagcc taggtcaagg aaagacagaa taatccagag 1560
ctgtggcaca catgacagac tcccagcagc cggggacctt gctgtcttct cgactcttcg 1620
ggcgtttctt tccatgtttg gctgttggtt ttagttttgg tgagccatgg gtgggccaga 1680
acatcactca actgcaattg ggctttcagg ttcttgccgg gagctctagg cactgggagg 1740
ctgtttcagg aaagtgagac tcaagaggaa gacagaaaag gttgtaacgt agaggaagtg 1800
agactggtga attggtttga tttttttcac atctagatgg ctgtcataaa gtttctagca 1860
tgttccccct cacctctccc cccccctgc cacttgaaac cttctactaa tcaagagaaa 1920
cttccaagcc aacggaatgg tcagatctca caggctgaga aattgttccc ctccaagcat 1980
ttcatgaaaa agctgcttct cattaaccat gcaaactctc acagcgatgt gaagagcttg 2040
acaagtcttt caaaataaaa agtaacaact tagaaacgg 2079

```

<210> 20
 <211> 14468
 <212> ADN
 <213> *Homo sapiens*

 <400> 20

tctaaggctg ttctgatctc ttcactctgc ccttcacctg gccctgtac ccccttcccc	60
tttggacccc ttgaaccctc ccaggacccc cgctcagccc ctteccgccc ccaaccgact	120
cttcccgaac gtcccttacc aaccgcgaga gcccctact gcgctttggc cacacccctt	180
acgcctcgct cccggccccc cctctgcccc tgaccgcgcc tgcgcaaggc gggcgcccta	240
aagtcctatt tcaactctgt gggaggagg ggaagggtg acgcaggcgc agtggcgtct	300
aaatttgggc ccactaatg cgtcggagca tctccgcgcc caggcggctc ctctcactg	360
cggcaacccg ggaaaactg tgaactaatc agaaaaagt gaaggcggga gatcttgggg	420
cgctgtccaa tggcgcggaa gagaacaaat gagctggcca atcggaacg gcacgggggc	480
gggctcgctc ggcgcgaagt tcgggcccgc gaattccgaa ggaggggtag gcgctgccc	540
cgcgcagagg ccgcgcccct cctggccccg gcttcttggc tgtcaaacag atgcagcaac	600
gtcggctcct gccgaggagc ccaaggggtc ccgggatccg ccgcacaggc tggcactgct	660
tgaagaggag gctactcgga gactgcgccg cgcgggtaga tccgaaacg ggctggggcg	720
gagtgggaaa aggcgggta tgccttgcac gatcgcgggg agctccttcc tgtttttatc	780
ccacctagag aagccgggaa gtaggggtt aggtccaatt tgttgagta ctttaaggact	840
cgtttgact ttcttttggg ggatgacagt ggattcattg ccctcggagg ttcaaccagt	900
tatgagttag ggattggcca gaagatcggg gcgcaggcaa gcaggagtgc tctattagga	960
taagcaagtt tgacaggaag aagctgctct tctccgaatt acacagagg gatgtgttcg	1020
tattgcacgt agacgttgt ataacaggac ctccctcccc gcgccccgc accccgacac	1080
acacaggagc tgccctaaagt atccttgctc tgcagattgg aggtcccca aatattttgt	1140
gatctgagga tccagctcaa gtgaggtgcc ataggacgtg ttcctgagtt tgcattgcac	1200
ggagaccttc ctggaatttt tcatttgcaa gtcggcttaa ccaattttgc attgagtcct	1260
aggctgcttg cactctgaat ttgggtatt caggtagtgt gctcaaagtt gaaaccgcat	1320
acagcacaac tcaagtttgc atcagactgg gaagcgaact taagccagcg gtgcgtggcc	1380
caggagtggg aaaggaaatg gatgcctgaa gtggaagagg tgggtgcagag ggggcaccgc	1440
ccatgctgcc ctgcttccaa ctgctgcgca tagggggcgg caggggcggg gatctctaca	1500
ccttccaccc cccgcgggg gctggctgca cctatcgctt gggccacagg gccgacctgt	1560
gtgatgtggc cctgcggccc cagcaggagc ctggcctcat ctctgggatc cacgccgaac	1620
tgcattgccg gccccggggt gatgactgga gggtcagcct ggaagaccac agcagccaag	1680
gtgagcatta agcagggcag ctttgcacct gggtggttga agcgccaggc tggaatgagt	1740

aaggtctcca caagaccctg ctgcctgcct cccatactcc catcagattg gatggatagt	1800
cgtgggccag accttcatct tcccaccaga agtgtgcaca gtcagaagct ctctgccaga	1860
ctgacccttt ttgggtccgt ttagctcata caggacctgg gatatcatca gaaagatata	1920
acagtgggga tgttctgagg ccactagagg ccaagtttag acttgattca gtttccagct	1980
ttgctgaggc actctgttcc tgggttaggg cagttctatg ttgaataatg tttttaataa	2040
tctgggcatg tctttctccg tgacttgagg cagttagcct cagaaagcct agattcacat	2100
ttgagttttg ccactgcctc ttggtaaagt cagctgtagg agtgttatgg ttattagact	2160
atagtagcca acattcatct agtgcttact gttatgagcc aggccctatt ttaagtgtat	2220
tgaatgtagg tgggtactaat attatcctca ttacagtaa aggaaaatga ggcacaaaga	2280
ggttaaggaa cttgtccagg gctgggcatg gtggtttaca cctataatcc agcactttgg	2340
gaggctaagg cagggtggat cacttgagct caggagtctg agaccagcct gggcaacatg	2400
gtgaaaacct gtctctacca aaaaattaat taatttttta aaaaaagcct gggcgcggtg	2460
gctcacgcct gtaatcccag cactttggga ggccgagatg ggcagatcac gaggtcagga	2520
gttcgagacc atcctgacca acatgttgaa accccatctg tgctgaaaaa aaaatacaaa	2580
aattagccag gtgtgtgggc gtgcacctgt aaccacagct actcaggagg ctgaagcagc	2640
agaatcactt gaaccggga ggcggagggt gcagtgaact gagatcgac cactgcactc	2700
cagcttgggc gacagagcga gactccatct caaacaaca aacaaacca aagcttgccc	2760
agggtcacat aactgtaag tggtagagct aggatctgaa cgagctggag ctgggggaga	2820
gtgagcatgt ttgaaaactg gaccttaggg cggggcacgg tggctcacgc ctgtaatccc	2880
agcactttgg gaggtgagg cgggcagatc aggaggtcag gagtatgaga ccagcctggc	2940
caacatggtg aaacctgtc tctgctaaaa ataaaaaat tagccagacg tgggtgcaca	3000
tgccctgtaat cccagctact caggaggctg aggcaggaga attgcttgaa cctggggaggc	3060
ggttcaagct tgggcaatag agcaaaactc catctcaaaa aaaaaaaaag aaagaaaaa	3120
aaagaagaaa gaaagaaaat tggaccttag gacagtgagg gcagggatcc tttgtaggaa	3180
agcacaagaa acacagactt gttcctagct gacaaggagt gtactgcctg gtacctgtca	3240
cctgctgagg ggcttaggat gtgagggaga atctgactac agtttcatat tcttcccag	3300
aatcataca gatttctcca ctctgactc tggtcatttc tgtttttgtc ctccatattt	3360
gcctggtgcc ccaccatcaa caggctactt ggtcaataat gtccgactcc caagaggcca	3420
caggctggaa ttgagtgatg gagacctcct gacctttggc cctgaagggc ccccaggaa	3480
cagccctcgt gagttctact tcatgttcca acaagtacga gtcaagcctc aggactttgc	3540
tgccattacc atcccacggt ctaggggaga agcccggtt ggggctggtt tccggcctat	3600

gctgcectcc cagggggctc cacagcggcc tctcagcacc ttctcccctg cccccaaggc 3660
cacactgatc ctaaaactcca taggcagcct cagcaagctc cggccccagc ccctcacctt 3720
ctcccctagt tggggtggac caaagagcct gcctgttccc gccccacctg gggaagtggg 3780
gaccacgcct tctgctccac cccaacgcaa tcggaggaaa tctgttcacc gagtgttggc 3840
ggaactggat gatgagagtg agcctcctga gaacccgcca ccggtcctta tggagcccag 3900
gaagaaactc cgtgtagaca aagccccact gactcccact gggtaagtgg agtcctcact 3960
tggccctctc agtgttttac tgcttttcga ttccctgtat ccctaggctg tgaggaggtc 4020
cccctgcctg gggggatggg caogggaggt ggaatagatg gaatggcaag acctgggtta 4080
gctctgatag gaaaagaaaa atatgtgcag gagaacatga gagtggggtt ggggcagtgc 4140
ttataaaaca accggagtga gcattgcctg ctttttacat tcatatggct ttaaccccat 4200
tcttctagtg cctaaggatg gggaactttc aggtcacac tagaggtttt taggcccacc 4260
ctatgtgttt ttaaggacag agtccaggct caccttagtt ctcagaccac tgtgcctctg 4320
tggcctcacc ctatgaccag ccattagggg gcaaggtcta ggcttctcc tacaggtttc 4380
cggtgacctt tgtgtctgtg tcaacttcctt cagaaatcga cgtggccgtc ctcggaagta 4440
cccagtgage gctcccatgg ctccccctgc agttgggggc ggggagccct gtgcagctcc 4500
ttgttgctgc ctgccccagg aagagacagt ggctgggtt cagtgtgatg gctgtgacgt 4560
ctgggtccat gtggcctgtg ttggctgcag catccaggct gccagggagg ccgacttccg 4620
atgcccaggg tgccgggctg gcattcagac ctaagggtcca ccgccaaggc accatcggac 4680
acacctgccc atgagtagac acagcagcga gcaaataggt ctgataaata ccccccttcc 4740
cttccctccc caagagggaa tgactacagg gaagaaggat ggattgatgt ggactcattc 4800
agggcctgga gcagacctg gtggccaaga cagaagagat ggtttcctgc caaagatatt 4860
gccacctcca ggaaattgcc agtgagctgg aagttccac tattacaagc cataaggcca 4920
tgttgccatg gacaccagaa tatctgtagt cagagcacct atcagttgca aaagccatgc 4980
ctgcaaccga tggaaaatgt aagagggagt tcttaaggtt cttggtggca tcaccaagg 5040
cattctggga aaacctaggg cctggcccca aaacttcctt actctgtggc tagtcctgct 5100
gccaacaaaa tcgtagcgac ctggcttttc acagctttgc ttttatttcc aagtcaagga 5160
caagccgctt cattcactcc tgggcattta ctcttctgtt gggctctgtg tattccttgc 5220
tttccaggga gaatgtgctt ggcaaggctt ggagaactaa ttcagaatct taggggaagg 5280
ggagagatgg aaatacaaac ctgcttactg gaaagggtca aatatatggg ttgagctgga 5340
ggtaggaata cagtaatta aggtttctag tttaaggga aacagatcta ttgccattta 5400
aataaggtaa ctgggatttg gttaagttca caaagatagc agaagattta ttacaggct 5460
tcacctgtac tgtcagggca agagaaagcc tggtaaacca gctacagcag tttaccagtg 5520

tgatggctgt gacacagctc cactccacgg gtggacacag cagagggcaa ctgggctggc	5580
ctggttcagt gtgaatcaaa ccgcttaacc cacacatggt acatgtgatt ttcttttctg	5640
agccttacac caagccaaac tattgtcaaa gcatcatttc tatagaaata aagccttacc	5700
ttgacctgtt ctattaaaac ctgccacatc cgccctttcc tacctagatt taatgagccc	5760
aagttttttt acatggaaga aatgactctg gggcaaagac ccctaataa ctagtggcag	5820
agccaggaat aaaacttgag taactaatga gtcacttatg ggcagagtat gcaaaaacct	5880
taagtggaaa ccaaatagac cctgggtatca agaaagcaca aagtattaat agaagtttct	5940
ggttgggggtg atctaggttc aacagaaata agatgatttc taagtataaa gccatttaag	6000
aattccagag taggggtggga aagcaaaaag ccagctctga acaggaaca gctacatggt	6060
gactgagtct atgggcacaa gtctctgcat cacaggcttt tgggaactag cctatcacag	6120
ggccctgtac aaataaactt ggctgcaatc ccagctctcc ctctgatgtt gtgtgacctt	6180
aaggagtgt aatggcacct tagtttcagg gtcactggg tatgagcatt ggatattccc	6240
atccccacct cagtaactga aggacaaaacc aagataagtg tgtctatcta ctgtgtccca	6300
agcttcttta tttaaagaaa aagtataca tgatgtggga ttaaaatcaa gagcatcatt	6360
gaacttcacc ttccctccaa ccagttgccc caaactcccc tgccccacc ctttgtgttc	6420
ccaattcctt ccttagtgaa tgaagaactt aatcccaaaa accctggcac aaactccagg	6480
ttttctttcc ctagtctct cctccccct gtccccatt cctagaaggg caggcacctc	6540
agtttgaatg catgggagag cccagagtgg tgacagagac agggggaaag gcttccccct	6600
cagggaagg gaccgaggag tacagtgcag tgaagtgagg gctcccatag cctgggttac	6660
caaatgggg ccctggggcc agaggaaagg aactgggtcc ccctgagaaa ggagaccag	6720
cagcctcaaa atcctctcgt tgtgcatagt cgctgcttga tcgcttgccc ttctggcgcc	6780
ggttacagaa ccacactcg accacctgcc agtgaatgac agaaaggaga atgacattag	6840
acaatgagct gagagacggg cctgactctg cttggacatt ctatccaaag ccaacagccc	6900
tagagcagtt agaggaggac attagagaat gagctgagac aggcctgact gcttgacat	6960
tctgtccaaa gccaacagcc ctagagcagt tggaggagcc agagctaggg aaagcgaggt	7020
ggtgacagg gaaagagatg gagcccgag agagacatgg cactcacatc cttctcgagc	7080
ccaagctgct gggcgatgtg gctgatctgc tgcagtgtgg gtttcgggca ctgcaggaac	7140
aaattctcca ggttgccct cactcgggtc tcgatactgg ttctgcttct ctttcgggcc	7200
tgcacgaggg tttctgcttt gcatatctgt gcaggtggga aggggtgac aagggaagc	7260
tttgacttg ctgagtaaca gcatcacagg ggtctgtgac tagatgtgtc agcagagcca	7320
ggtggtggtg tgaagggca ggatcctgga aggggtggct ctggacctta tcccagcaga	7380

actgaggaat	ttcactccat	cccactgaga	accactgcac	caaagacgga	gagctacgag	7440
ccagtgatgg	aagcaatgga	aattaggcca	agaaagggaa	ggtccccggg	tatccccctc	7500
ccaccottac	ctcctgaaga	ttttcattgt	tgtcagcttc	ctccaccac	ttctgcagca	7560
agggcgcag	cttacacatg	ttcttgaagc	taagctgcag	agcctcaaag	cggcagatgg	7620
tcgtttggct	gaataccttc	cctgggggag	gccagtcaaa	agagaagcaa	aatgagggag	7680
cacgcagggc	cotttgtagc	ctgagatcca	agcttaccac	ctcttcccag	agggagctca	7740
aagcccaagc	atcttctccc	tctccctact	cctcttcctg	ggtgagggtg	gagtctgccc	7800
ctgcccctcc	ccactaggtt	cagggatact	ccttagaggg	gagatgcggg	cagaatctgc	7860
agaggggaac	ccaccaata	gaacccccag	ggtgagcccc	acatcggcct	gtgtatatcc	7920
cagggtgatc	ctcttctgct	tcaggagctt	ggcaaattgc	tcgagttctt	tctgcagagc	7980
tttgatgtcc	tgggactgga	ttttaaaagg	cagaagactt	gtaagaacat	aaacacacca	8040
gttatcaatc	tcccctttcc	attcgggatt	caagaaccta	cgtgtggccc	caaggaatag	8100
tctgtagaag	tgcattctgc	ttccaagctg	cccacctaac	ttctagaaat	aacctacca	8160
caaatgtcat	tcacccattc	cctgttcact	gactcatgca	tgtacaaaag	gactactctt	8220
ccccagaaa	ctggcacatc	caagggatgc	agagcatggg	gaaaggacag	aaagagagac	8280
cctggcctcg	aggagaacac	ctgtcaggtt	atgaagggtt	gaagttcttt	gctgggcgcg	8340
gtggctcatg	cctataattc	cagcactttg	ggaggccgag	gtgggcagat	cacgaggtca	8400
ggagttcaag	accagcatgg	ccaacatggg	gaaaccccg	ctctactaaa	aacacaaaaa	8460
ttagctgggc	acgggtggac	gcacctgtaa	tcccagctac	tcaggaggct	gaggcaggag	8520
aatcacttga	acccgggagg	cggaggttgc	agtgcagctg	gatcacgcca	ctgcactcca	8580
gcctgggtga	cagagcaaga	ctctgtctca	aagaaaaaaa	aaaagaagat	agttcattta	8640
atacctgcaa	aattctctca	ctcaagtatc	acccccagtt	taaggatgtt	ttgagattag	8700
agaaatagat	aagctgctaa	gttctgggtt	aattaaaaag	gaagagcatc	atgtctcaga	8760
agctaaattc	agtatatact	ctccccagct	tgctttgagg	gtcccacaaa	ctataacatg	8820
gcatgcatac	acacaaacac	agcaaaaaag	taacaggtgt	cataagaatg	gataaagtgc	8880
tttgtgtgta	cttactcctc	atTTTTTaaa	ttgattatcc	ctcatcttta	ctgtatcttt	8940
ttcactatag	aggcatccta	attgattttt	aaattcaaga	gatttatcga	gcaccttcta	9000
taagccagcg	gctatacaaa	gtggacaaag	agccctgaca	tccagcatga	cagaagtgtc	9060
attcggcact	tgttcttcaa	gttgccact	tggatctctt	ccaagtgcac	tttctttttt	9120
tccctgcctt	ataacttttt	aataataaac	ttccactcct	gctctgaaaa	ataaaaaagt	9180
aaataaaata	aaaaatggcc	aggcacagtg	gctcatgtct	gtaaatccta	gcactttggg	9240
aggccaaggt	gggcagactg	cttgagccca	agagttagaa	agcagcctgg	gtaacatagt	9300

gagacccgtg ccgccccctt tcccacccct gctgcctcta tttaaaaaat atatatatat 9360
 tatggaaaaa agcaaagcag tccgggcgca gtggtcatgc ctgtaatccc ttcactttgg 9420
 gaggccaagg tgggtagatc acttgaggtc aggagttcaa gactagcctg gtcaacatag 9480
 tgagactctg tctctactaa aaatacaaaa attagctggg catcatggcg ctcccctata 9540
 atcccagcta ctccaggaggc tggggcagga gaattgcttg aacctaggag gtggagtttg 9600
 cagtgagcca agatgcacc actgcactcc agcctgaggg acagagttag actccatctc 9660
 aaaaattaaa aaaaaataa agcagtctat aggagttagg taaaggaggg aaggagatta 9720
 tggaggaggg tgacactttt aaagacagag aagggtgatt tttgagcaaa ggacaagagt 9780
 ctaatgtggc aaggccctga agtgggcctt ccagagccca aagctggtct ggtggctagg 9840
 tagatcctgt tgcagacata gtgactttgt tttagtccaa gtgaaatgat ctctcaccct 9900
 ttttctcccc cccaagacg gaatctcgtt ctatcgccca ggctggagtg ctgtggcgtg 9960
 atcttggtc actgcaatct ccgcctctg ggttcaagct attctgcctc agccgcctga 10020
 gtagctggga ctacaggcac ccaccacat gcccggttaa ttttgtatt tttagtagat 10080
 atggggtttc accatgttgg ccaggctggt caggagacct caagtgatct gtccaccttg 10140
 gcttccaaa gtgctgggat tacaggtgtg aaccaccgca cctagcctca cctttttttt 10200
 ttttttttg agagtttcgc ttttgtgccc taggctggag tgcactggcg cgatctcggc 10260
 tcaccgcaac ctacatctcc caggttcaag cgattctcct gctcagctt cctgagtagc 10320
 tgagattaca ggcattgcgc accacgcccc gctaattttg tatttttagt agagatgggg 10380
 tttcgccatg ttggtcaggc tggactcgaa ctcccaacct caggtgattc gcctgcctcg 10440
 gcctccaaa gtgcctggcc acacctttta aaacactgac tctagttgac gtgttgcca 10500
 cagacagtag ggaggaagca gtataatttg agaagctact gcggaatcc cagcagagat 10560
 gatggtggct gaggccaggg ttaggttgtg attgattcag gatgtttctt aaggatagga 10620
 tgtaggacgt gaaagaaact gaggatgact gggtttgccc ttgagcaact gggatgatcag 10680
 ggtggagcag ttcagggagc catcacaaga gacagaaaac gcggtagtca tctggtgtct 10740
 aaatggcatt taagccttga ggggtgggtga gaggaaggaa gggtagatag agcagaggtt 10800
 gaaggactga gccctggggc atgcatatg aggctgccgg cggacagagg tgcacagcta 10860
 gtgagaaaaa aacaaggcct tttttagtgc ctgaagcctc aaggaaagtgt ttcaatggtg 10920
 cttgatcata tcaatttcaa ataggctgtt ttcacccca acttctgctc agccaataac 10980
 tcaaaactgat aaatgccctc tgctatcctg gattttccaa attctgtttt ggggttttgg 11040
 aataaacact ggtccaaatc ctgccttcat catttagcag ttaaaaccgg ttaaatagga 11100
 taataatacc tcccctagg agattttgtg ctggttaatg agataatgat gtataaacgg 11160

agcacacagc caggcactta ggaagtggac cacaattgcc agccattatc attcaaggct 11220
 cagcagtgc ctcctgcgaa gaggttgggg cttctcggtc actccagaaa ccagtccacac 11280
 ctttctgtga ggtctcaagg cttagtattt aatctetaat tgcttacact tgctgccttg 11340
 gaggactgga agatacatct ttaatagtcc tcagcagggc tggatgcctt caatcccgca 11400
 gcagctctat atttgcaaat ggctgggaga aatctctcac catttttctt gtttacaact 11460
 ttggaactga ggctgaagtc aatcaaaatc cagctttcta caaggggtgc caggggtgtgc 11520
 aocctaacac agtggccagt cattggcctg aggcagagat cgggggaaga caagccctat 11580
 acttgactgg aggtaaaccc agctcacaac ggcacacac acagcccaa caggagatcc 11640
 tatcagaaac gagtccacac ctagactttc aggaacaata atcctggaat gagcactgtt 11700
 tttaccctca ggctatgctt aaccctaagg ccaaaatctt gggctctgata aggggtcaat 11760
 tttcaagcag gactaagggt gggaaaagg gctcaaacca accccaagct gggctctggtg 11820
 ctggggccagt aatgagtgc cagaccctgg gcaggcctag gagatgtgag agaccctgac 11880
 aagggctggg ccagacagag caaaggccag cctggggccag cttccgactc tcccaggccg 11940
 ctctgccctc acctgcagtt gtctcttcga aatccagctt ccagttccca cctggccctt 12000
 gcctgccagg gctgcctgca gttgatacac accctccct ggccagggea gctgaccctg 12060
 cctgctcctc tcctgggtgc caggtctggg cagctgcagg tgaccacttc cccatcaggc 12120
 tgccctgtca tgaccactc cccacacccc aaccccgctg aagctcactt gcctcctccg 12180
 ggttttgtc cagcttctcc ttctccagct tcacggcacc aggggtgacg gtgcagggt 12240
 ccggggaggc cccatcggag ttgctctcca ccccgactcc tgcttcgccc tcaggctgag 12300
 aggtctccaa gccgccttg ggcaactagcc ccaactccaac ctggggccca cagtacgcca 12360
 tccccccaca gaactcatc ggcggggggc atgggggaat ccccccacac tcagagcctg 12420
 gcccaccccc cggcccgatt cctggccctc caggagggcc ttggaagctt agccaggctc 12480
 gaggatcaac ccagcccggc tccggccccc ctggcccatc acctccacca cctggagggg 12540
 gcgagaaggc aaaatctgaa gccagggtgc ccgccatggg gaagggaagg gcccgaaggc 12600
 gggggcctgg tgaaatgagg gcttgccaag ggactactca accctctctt ccctcccag 12660
 tcccacccac tagccttgac ctctggcccc gccccctgga tgggtggagg agagggaggt 12720
 ggggggagaa actgaggcga aggatgtttg cctaattggtg gtggcaatgg tgtctgtgga 12780
 aggggaaaac cgggagacac aactggcgcc cctccaggac ctcagtgcag gtccccaca 12840
 gaaacttttt ttatttttat tttttaagac agggctctac tttgttgccc agactggagt 12900
 gcagtggagt acaatgatgg ctcaatgtag cctcgatcta ctgggcaaaa gcaatccttc 12960
 tgctccagcc tcctaagtgg ctgggactac aggcttggac cactgtgccc tgtagtttt 13020
 tttattttta gtagagatgg ggcttgcta tgttaccag gctggtcttg aattcctgtc 13080

ctcaagaaat cctcccgccct ctgccgcccga gtgtcatgat taaaggcgtg agccaccaca 13140
 cccaactttc aactcccaac ccgctccctg gcaactctctc aggcctctgca catcccagct 13200
 gtctggaatc actcccacac ctccatgttc ttcaggaacc cagggtgcttg accccctctc 13260
 cacagacctc tggcactgtg ccttcagggg ccagtcaccc tctcagctcc tcaaatttat 13320
 tgaatgtgtg tgtggcgcta tccctcaatg catcaacagc cataagcaca atggccagct 13380
 gctcccttat gccttcccc gatccatcca gaatcctagg cattcccatc ccgatactgg 13440
 ccaaatccag ccaccccgca gcctgggtgc ctggcaccat ctgccagcc tgccaaattt 13500
 caccatct tcaagagtag actgccagac aaggcctccg tgctatatcc cccaccccc 13560
 catccccca cccctccgtc ttccagaatc agactccaga ctctcctcat ctaacagact 13620
 aagggttg cccctacttc cccttcaagg gaccagactt tggactgatt gggcctcagt 13680
 ttcccaacct ttgctgaaac agagtataa gacacccgct ttgggcccccc tccactatgg 13740
 aacctgcaca tcaggttcct tgctccctc tcaaccacaaa ctcagacatc taataccacg 13800
 gtaggccccg ttctccctcc cccacctccc tggcccaggc ctccagccct aggccttggg 13860
 tggggaaaac caggggtgtg ggggtgtgga gaaaaatat ctgacttcag gttcaaagaa 13920
 gcctgggagg gactggggga agggggcagg acaatggcct tggctggaca atcccggtcc 13980
 ccagaggggg cagctctaac cctaaacaag tgctcaacc ttgaatggg ctggatggct 14040
 cccctgggga ctgcttcctg ctccccaccc cccagtcct aatccctca cacagaatcc 14100
 ccttcagaga cgctaaaagg agctccagca accccctct gcaatccct caaagactga 14160
 gcctcagacg ggcaccaagg gccccctaca gggacctagg tatctagtcc ctcttctc 14220
 tgggggactc aggcgtccag cttcatcgtg catccctccc cgagcccgga agattgaggg 14280
 atgtgctttg tttagtgggg ctggctggca gaaagacgca gaggaggtgg cgagtgattt 14340
 gtggaggggt gcaggaaggc tgccctaagc tccccttcag ggtctgttt tctgggcctg 14400
 gcctgagtat cctgaggctc atgctgctgg tctagtgtt gattctgtt gcaagagaat 14460
 agccaacg 14468

<210> 21
 <211> 3021
 <212> ADN
 <213> *Homo sapiens*
 <400> 21

tcgtctcctg ccgaggagcc caaggggtcc cgggatccgc cgcacaggct ggcactgctt 60
 gaagaggagg ctactcggag actgcgcgc gcggattgga ggctcccaa atattttgcg 120
 atctgaggat ccagctcaag tgaggtgcc taggacgtgt tcctgagttt gcattgcacg 180

gagaccttcc tgggaattttt catttgcaag tcggcttaac caatthttgca ttgagtecta 240
 ggctgcttgc actctgaatt tgggctattc aggtagtgtg ctcaaagttg aaaccgcata 300
 cagcacaact caagtttgca tcagactggg aagcgaactt aagccagcgg tgcgtggccc 360
 aggagtggga aaggaaatgg atgcctgaag tggagaggt ggtgcagagg gggcacccgcc 420
 catgctgccc tgcttccaac tgctgcgcac agggggcggc agggggcggg atctctacac 480
 ctccaccccc cccgcggggg ctggctgcac ctatcgcttg ggccacaggg ccgacctgtg 540
 tgatgtggcc ctgcggcccc agcaggagcc tggcctcatc tctgggatcc acgccgaact 600
 gcatgccgag ccccggggtg atgactggag ggtcagcctg gaagaccaca gcagccaagg 660
 tactttggtc aataatgtcc gactcccaag aggtcacagg ctggaattga gtgatggaga 720
 cctcctgacc tttggccctg aagggccccc aggaaccagc ccctcggagt tctacttcat 780
 gttccaacaa gtacgagtca agcctcagga ctttgctgcc attaccatcc cacggtctag 840
 gggagaagcc cgggttgggg ctggtttccg gcctatgtg ccctcccagg gggctccaca 900
 gcggcctctc agcaccttct cccctgcccc caaggccaca ctgaccta aactccatagg 960
 cagcctcagc aagctccggc cccagccctt cactttctcc cctagtggg gtggacccaa 1020
 gagcctgctt gttcccgccc cacttgggga agtggggacc acgccttctg ctccacccca 1080
 acgcaatcgg aggaatctg ttcaccgagt gttggcggaa ctggatgatg agagtgagcc 1140
 tcttgagaac ccgccaccgg tccttatgga gccaggaag aaactccgtg tagacaaagc 1200
 cccactgact cccactggaa atcgacgtgg ccgtcctcgg aagtaccag tgagcgctcc 1260
 catggctccc cctgcagttg ggggcgggga gccctgtgca gctccttgtt gctgcctgcc 1320
 ccaggaagag acagtggcct ggggttcagt tgatggctgt gacgtctggt tccatgtggc 1380
 ctgtgttgcc tgcagcatcc aggtgccag ggaggccgac ttccgatgcc cagggtgccg 1440
 ggctggcatt cagacctaag gtccaccgcc aaggcaccat cggacacacc tgcccatgag 1500
 tagacacagc agcgagcaaa taggtctgat aaataccccc ctcccttcc ctcccagga 1560
 gggaatgact acaggaaga aggatggatt gatgtggact cattcagggc ctggagcaga 1620
 ccctgggtggc caagacagaa gagatggttt cctgccaaag atattgccac ctccaggaaa 1680
 ttgccagtga gctggaagtt cccactatta caagccataa ggccatgttg ccatggacac 1740
 cagaatatct gtagtcagag cacctatcag ttgcaaaagc catgcctgca accgatggaa 1800
 aatgtaagag ggagttctta aggttcttga tggcatcacc caaggcattc tgggaaaacc 1860
 tagggcctgg ccccaaaact tccctactct gtggctagtc ctgctgcaa caaaatcgta 1920
 gcgacctggc ttttcacagc tttgctttta tttccaagtc aaggacaagc cgcttcattc 1980
 actcctgggc atttactctt cttgtgggtc tgtgatattc cttgctttcc agggagaatg 2040
 tgcttgcaa ggtctggaga actaattcag aatcttaggg gaaggggaga gatggaaata 2100

```

caaacctgct tactggaaag gtgcaaatat atggggttgag ctggaggtag gaatacaggt 2160
aattaaggtt tctggtttaa gggaaaacag atctattgcc atttaaataa ggtaactggg 2220
atttggttaa gttcacaaag ttagcagaag atttatttac aggcttcacc tgtactgtca 2280
gggcaagaga aagcctggta aaccagctac agcagtttac cagtgtgatg gctgtgacac 2340
agctccactc cacgggtgga cacagcagag ggcaactggg ctggcctggg tcagtgtgaa 2400
tcaaaccgct taaccacacac atggtacatg tgattttctt ttgtgagcct tacaccaagc 2460
caaaactattg tcaaagcacc atttctatag aaataaagcc ttatcttgac ctgttctatt 2520
aaaacctgcc acatccgccc tttcctacct agattttaatg agcccaagtt tttttacatg 2580
gaagaaatga ctctggggca aagaccoccta atgaactagt ggcagagcca ggaataaaac 2640
ttgagtaact aatgagtcac ttatgggcag agtatgcaaa aaccttaagt ggaaccaaaa 2700
tagaccctgg tatcaagaaa gcacaaagta ttaatagaag tttctggttg gggatgacta 2760
ggttcaacag aaataagatg atttctaagt ataaagccat tttagaattc cagagtaggg 2820
tgggaaagca aaaagccagc tctgaacagg taacagctac atggtgactg agtctatggg 2880
caaaagttct tgcatacacag gcttttggga actagcctat cacaggggcc tgtacaaata 2940
aacttggtcg caatcccaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3000
aaaaaaaaa aaaaaaaaaa a 3021

```

<210> 22
 <211> 1080
 <212> ADN
 <213> *Homo sapiens*

<400> 22

```

atgctgccct gcttccaact gctgcgcata gggggcggca ggggcggtga tctctacacc 60
ttccaccccc ccgcccgggt tggctgcacc tategcttg gccacagggc cgacctgtgt 120
gatgtggccc tgcggcccca gcaggagcct ggccctcatc ctgggatcca cgccgaactg 180
catgccgagc cccgggggtg tgactggagg gtcagcctgg aagaccacag cagccaaggt 240
actttggtca ataattgtccg actcccaaga ggtcacaggc tggaattgag tgatggagac 300
ctctgacct ttggcctga agggcccca ggaaccagcc cctcgagtt ctacttcacg 360
ttccaacaag tacgagtaa gcctcaggac tttgctgcca ttaccatccc acggtctagg 420
ggagaagccc ggggtggggc tggtttcggg cctatgctgc cctcccaggg ggctccacag 480
cggcctctca gcacctctc cctgcccc aaggccacac tgatcctaaa ctccataggg 540
agcctcagca agctccggcc ccagcccctc accttctccc ctagttgggg tggaccaaag 600

```

```

agcctgcctg ttcccgcgcc acctggggaa gtggggacca cgccttctgc tccaccccaa 660
cgcaatcgga ggaaatctgt tcaccgagtg ttggcggaac tggatgatga gagtgagcct 720
cctgagaacc cgccaccggt ccttatggag cccaggaaga aactccgtgt agacaaagcc 780
ccactgactc ccactggaaa tgcacgtggc cgtcctcgga agtaccagtg gagcgctccc 840
atggctcccc ctgcagttgg ggcggggagc cctgtgcagc tccttgttgc tgcctgcccc 900
aggaagagac agtggcctgg gttcagtggt atggctgtga cgtctggttc catgtggcct 960
gtgttggctg cagcatccag gctgccaggg aggccgactt ccgatgcca gggtgccggg 1020
ctggcattca gcctaaggtc caccgccaag gcaccatcgg acacacctgc ccatgagtag 1080

```

<210> 23
 <211> 4651
 <212> ADN
 <213> *Homo sapiens*

5

<400> 23

```

ccgccacccc gacacacaca ggagctgcct aaagtatcct tgccttgcaag attggaggct 60
ccccaaatat tttgtgatct gaggatccag ctcaagttag gtgccatagg acgtgttcct 120
gagtttgcat tgcacggaga ccttcctgga atttttcatt tgcaagtcgg cttaaccaat 180
tttgcaattga gtcctaggct gcttgcactc tgaatttggg ctattcaggt agtgtgctca 240
aagttgaaac cgcatacagc acaactcaag tttgcatcag actgggaagc gaacttaagc 300
cagcgggtgcg tggcccagga gtgggaaagg aaatggatgc ctgaagtgga agaggtggtg 360
cagagggggc accgcccatt ctgccctgct tccaactgct gcgcataagg ggcggcaggg 420
gcggtgatct ctacaccttc ccccccccg ccggggctgg ctgcacctat cgcttgggcc 480
acagggccga cctgtgtgat gtggccctgc ggcccagca ggagcctggc ctcatctctg 540
ggatccacgc cgaactgcat gccgagcccc ggggtgatga ctggagggtc agcctggaag 600
accacagcag ccaaggtgag cattaagcag ggcagctttg cccctgggtg gttgaagcgc 660
caggctggaa tgagtaaggc ctccacaaga ccctgctgcc tgcctccat actcccatca 720
gattggatgg atagtcgtgg tccagacctt catcttccca ccagaagtgt gcacagttag 780
aagctctctg ccagactgac cctttttggt ccggttttag tcatacagga cctgggatat 840
catcagaaaag atatcacagt ggggatgttc tgaggccact agaggccaag tttagacttg 900
attcagtttc cagctttgct gaggcactct gttcctgggt tagggcagtt ctatgttgaa 960
taatgttttt aataatctgg gcatgtcttt ctccgtgact tgaggcagtt agcctcagaa 1020
agcctagatt cacatttgag ttttgccact gcctcttggg aaagtcagct gtaggagtg 1080
tatggttatt agactatagt agccaacatt catctagtgc ttactgttat gagccaggcc 1140

```

10

ES 2 565 496 T3

ctattttaag	tgtattgaat	gtaggtggta	ctaataattat	cctcatttac	agtaaaggaa	1200
aatgaggcac	aaagaggta	aggaacttgt	ccagggctgg	gcatggtggt	ttacacctat	1260
aatccagcac	tttgggaggc	taaggcaggg	tggatcactt	gagctcagga	gttcgagacc	1320
agcctgggca	acatggtgaa	aacctgtctc	tacaaaaaaa	ttaattaatt	ttttaaaaaa	1380
agcctgggcg	cggtggctca	cgcctgtaat	cccagcactt	tgggaggccg	agatgggcag	1440
atcacgaggt	caggagtctg	agaccatcct	gaccaacatg	ttgaaacccc	atctgtgctg	1500
aaaaaaaaat	acaaaaatta	gccaggtgtg	gtggcggtgca	cctgtaaccc	cagctactca	1560
ggaggctgaa	gcagcagaat	cacttgaacc	cgggaggcgg	aggttgaggt	gagctgagat	1620
cgcaccactg	cactccagct	tgggcgacag	agcgagactc	catctcaaac	aaacaaacaa	1680
acccaaaagct	tgcccagggt	cacataactg	gtaagtggta	gagctaggat	ctgaacgagc	1740
tggagctggg	ggagagttag	catgtttgaa	aactggacct	tagggcgggg	cacggtggct	1800
cacgcctgta	atcccagcac	tttgggaggc	tgaggcgggc	agatcaggag	gtcaggagta	1860
tgagaccagc	ctggccaaca	tggtaaaacc	ctgtctctgc	taaaaataaa	aaaattagcc	1920
agacgtggtg	gcacatgcct	gtaatccag	ctactcagga	ggctgaggca	ggagaattgc	1980
ttgaacctgg	gagggcgttc	aagcttgggc	aatagagcaa	aactccatct	caaaaaaaaaa	2040
aaaaagaaag	aaaaaaaaag	aagaaagaaa	gaaaattgga	ccttaggaca	gtgagggcag	2100
ggatcctttg	taggaaagca	caagaaacac	agacttggtc	ctagctgaca	aggagtgtac	2160
tgectggtac	ctgtcacctg	ctgaggggct	taggatgtga	gggagaatct	gactacagtt	2220
tcatattctt	ccccagaaat	catacagatt	tctccactcc	tgactctggt	catttctgtt	2280
tttgtcctcc	atatttgcc	ggtgccccac	catcaacagg	tactttggtc	aataatgtcc	2340
gactcccaag	aggtcacagg	ctggaattga	gtgatggaga	cctcctgacc	tttggccctg	2400
aagggccccc	aggaaccagc	ccctcggagt	tctacttcat	gttccaacaa	gtacgagtca	2460
agcctcagga	ctttgctgcc	attaccatcc	cacggtctag	gggagaagcc	cgggttgggg	2520
ctggtttccg	gcctatgctg	ccctcccagg	gggctccaca	gcggcctctc	agcaccctct	2580
ccctgcccc	caaggccaca	ctgacccata	actccatagg	cagcctcagc	aagctccggc	2640
cccagccct	caccttctcc	cctagttggg	gtggacccaa	gagcctgcct	gttcccgccc	2700
cacctgggga	agtggggacc	acgccttctg	ctccacccca	acgcaatcgg	aggaaatctg	2760
ttcaccgagt	gttggcggaa	ctggatgatg	agagttagcc	tcctgagaac	ccgccaccgg	2820
tccttatgga	gcccaggaag	aaactccgtg	tagacaaagc	cccactgact	cccactgggt	2880
aagtggagtc	ctcacttggc	cctctcagtg	ttttactgct	tttcgattcc	ttgtatccct	2940
aggctgtgag	gaggtcccc	tgctggggg	gatgggcacg	ggaggtggaa	tagatggaat	3000
ggcaagacct	gggttagctc	tgataggaaa	agaaaaatat	gtgcaggaga	acatgagagg	3060

tgggggtgggg	cagtgccttat	aaaacaaccg	gagtgcagcat	gtcctgccttt	ttacattcat	3120
atggctttaa	ccccattcct	ctagtgccta	aggatgggga	actttcaggc	tcacactaga	3180
ggtttttagg	cccaccctat	gtgtttttaa	ggacagagtc	caggctcacc	ttagtctca	3240
gaccactgtg	cctctgtggc	ctcacccctat	gaccagccat	agggtggcaa	ggtctaggcc	3300
ttctcctaca	ggtttcgggt	gacccttggt	tctgtgtcac	ttcagaaatc	gacgtggccg	3360
tcctcggaag	taccagtgga	gcgctcccat	ggctccccct	gcagttgggg	gcggggagcc	3420
ctgtgcagct	ccttggtgct	gcctgccccca	ggaagagaca	gtggcctggg	ttcagtgga	3480
tggctgtgac	gtctggttcc	atgtggcctg	tggtggctgc	agcatccagg	ctgccaggga	3540
ggccgacttc	cgatgccag	ggtgccgggc	tggcattcag	acctaaggtc	caccgccaag	3600
gcaccatcgg	acacacctgc	ccatgagtag	acacagcagc	gagcaaatac	gtctgataaa	3660
tacccccctt	cccttccctc	cccaagaggg	aatgactaca	gggaagaagg	atggattgat	3720
gtggactcat	tcagggcctg	gagcagaccc	tggtggccaa	gacagaagag	atggtttcct	3780
gccaaagata	ttgccacctc	caggaaattg	ccagtgcgct	ggaagttccc	actattacaa	3840
gccataaggc	catgttgcca	tggacaccag	aatatctgta	gtcagagcac	ctatcagttg	3900
caaaagccat	gcctgcaacc	gatggaaaat	gtaagaggga	gttcttaagg	ttcttggtgg	3960
catcacccaa	ggcattctgg	gaaaacctag	ggcctggccc	caaaacttcc	ctactctgtg	4020
gctagtcctg	ctgccaacaa	aatcgtagcg	acctggcttt	tcacagcttt	gcttttattt	4080
ccaagtcaag	gacaagccgc	ttcatttact	cctgggcatt	tactcttctt	gtgggtctgt	4140
gatattcctt	gctttccagg	gagaatgtgc	ttggcaagg	ctggagaact	aattcagaat	4200
cttaggggaa	ggggagagat	ggaaatacaa	acctgcttac	tggaaagggt	caaataatg	4260
ggttgagctg	gaggtaggaa	tacaggtaat	taaggtttct	agtttaagg	aaaacagatc	4320
tattgccatt	taaataagg	aactgggatt	tggttaagtt	cacaaagata	gcagaagatt	4380
tatttacagg	cttcacctgt	actgtcagg	caagagaaag	cctggtaaac	cagctacagc	4440
agtttaccag	tgtgatggct	gtgacacagc	tccactccac	gggtggacac	agcagagggc	4500
aactgggctg	gcctggttca	gtgtgaatca	aaccgcttaa	cccacacatg	gtacatgtga	4560
ttttcttttg	tgagccttac	accaagccaa	actattgtca	aagcatcatt	tctatagaaa	4620
taaagcctta	tcttgacctg	ttctattaaa	a			4651

<210> 24
 <211> 5522
 <212> ADN
 <213> *Homo sapiens*

<400> 24

gaattccgaa ggaggggtag ggcgtgcccg cgcgcagagg ccgcgccct cctggccccg	60
gcttcttggc tgtcaaacag tagcagcaac gtccgctcct gccgaggagc ccaaggggtc	120
ccgggatccg ccgcacaggc tggcactgct tgaagaggag gctactcgga gactgcgccg	180
cgccgggtaga tccgaaacgg ggctggggcg gagtgggaaa aggcggggtg tgccttgcat	240
gacgcggggg agctccttcc tgtttttatc ccacctagag aagccgggaa gtaggggttt	300
aggtccaatt tgttgagta cttaggact cgtttgcact ttcttttggg ggatgacagt	360
ggattcattg ccctcgagg ttcaaccagt tatgagttag ggattggcca gaagatcggg	420
gcgcaggcaa gcaggagtgc tctattagga taagcaagtt tgacaggaag aagctgctct	480
tctccgaatt acacagaggt gatgtgttcg tattgcacgt agacgtgtgt ataacaggac	540
ctccttcccc gcgccccgcc accccgacac acacaggagc tgcctaaagt atccttgctt	600
tgaagattgg aggtcccca aatatttggg gatctgagga tccagctcaa gtgaggtgcc	660
ataggacgtg ttctgagtt tgcattgcac ggagaccttc ctggaatttt tcatttgcaa	720
gtccgcttaa ccaatttggc attgagtcct aggtgcttg cactctgaat ttgggctatt	780
caggtagtgt gctcaaagtt gagaccgat acagaacaac tcaagtttgc atcagactgg	840
gaagcgaaact taagccagcg gtgcgtggcc caggagtggg aaaggaaatg gatgcctgaa	900
gtggaagagg tgggtgcagag ggggcacccg ccatgctgcc ctgcttccaa ctgctgcgca	960
tagggggcgg caggggcggg gatctctaca ccttcacccc cccgcggggg gttggctgca	1020
cctatcgctt gggccacagg gccgacctgt gtgatgtggc cctgcggccc cagcaggagc	1080
ctggcctcat ctctgggatc cagccgaac tgcattccga gccccggggg gatgactgga	1140
gggtcagcct ggaagaccac agcagccaag gtgagcatta agcagggcag ctttgccctt	1200
gggtgggtga agcgcagggc tggaatgagt aaggtctcca caagaccttg ctgtctgctt	1260
cccatactcc catcagattg gatggatagt cgtgggccag accttcattt tcccaccaga	1320
agtgtgcaca gtcagaagct ctctgccaga ctgacctttt ttggtcccg ttagctcata	1380
caggacctgg gatatcatca gaaagatct acagtgggga tgttctgagg ccactagagg	1440
ccaagtttag acttgattca gtttccagct ttgctgaggc actctgttcc tgggttaggg	1500
cagttctatg ttgaataatg tttttaataa tctgggcatg tctttctccg tgacttgagg	1560
cagttagcct cagaaagcct agattcacat ttgagtttgg ccactgcctc ttggtaaagt	1620
cagctgtagg agtggtatgg ttattagact atagtagcca acattcatct agtgcttaact	1680
gttatgagcc aggcctatt ttaagtgtat tgaatgtagg tggtaactaat attatctca	1740
tttacagaaa aggaaaatga ggcacaaaga ggttaaggaa cttgtccagg gctgggcatg	1800
gtggtttaca cctataatcc agcacttttg gaggctaagg cagggtggat cacttgagct	1860
caggagttcg agaccagcct gggcaacatg gtgaaaacct gtctctacca aaaaattaat	1920

tagtTTTTta	aaaaaagcct	gggcgcgggtg	ggtcacgcct	gtaatcccag	cactttggga	1980
ggccgagatg	ggcagatcac	gaggtcagga	gttcgagacc	atcctgacca	acatgttgaa	2040
accccatctg	tgctgaaaaa	aaaaaaccca	aattaaccag	gtgtggtggc	gtgcacctgt	2100
aaccccagct	actcaggagg	ctgaagcagc	agaatcactt	gaaccggga	ggcggagggt	2160
ggagtgagct	gagatcgcac	cactgcactc	cagcttgggc	gacagagcga	gactccatct	2220
caaacaaaca	aacaaaccaa	aagcttgccc	agggtcacat	aactggtaag	tggtagagct	2280
aggtactgaa	cgagctggag	ctgggggaga	gtgagcatgt	ttgaaaactg	gaccttaggg	2340
cggggcacgg	tccgtcacgc	ctgtaatccc	agcactttgg	gaggctgagg	cgggcagatc	2400
aggaggtcag	gagtatgaga	ccagcctggc	caacatggta	aaaccctgtc	tctcgtaaaa	2460
ataaaaaaat	tagccagacg	tggtggcaca	tgctgtaat	cccagctact	caggaggctg	2520
aggcaggaga	attgcttgaa	cctgggaggg	ggagtgcagt	gagctgagat	tgccactactg	2580
cactccagct	tgggcaatag	agcaaactc	catctcaaaa	aaaaaaaaaa	gaaagaaaaa	2640
aaaagaagaa	agaaagaatt	ttggccctta	ggacagttag	ggcagggttc	ctttgtggga	2700
aagcacaaga	aacacagatt	tgttcctagc	tgacaaggag	tgtactgcct	ggtacctgtc	2760
acctgctgag	gggcttagga	tgtgaggagg	aatctgacca	cagtttcata	ttcttcccca	2820
gaaatcatac	agatttctcc	actcctgact	ctgggtcaatt	ctgtttttgt	cctccatatt	2880
tgcttggtgc	cccaccatca	acaggtactt	tggtcaataa	tgtccgactc	ccaagaggtc	2940
acaggctgga	attgagtgat	ggagacctcc	tgacctttgg	ccctgaaggg	ccccaggaa	3000
ccagcccctc	ggagttctac	ttcatgttcc	aacaagtacg	agtcaagcct	caggactttg	3060
ctgccattac	catcccacgg	tctaggggag	aagcccgggt	tggggctggg	ttccggccta	3120
tgctgcctc	ccagggggct	ccacagcggc	ctctcagcac	cttctcccct	gcccccaagg	3180
ccacactgat	cctaaactcc	ataggcagcc	tcagcaagct	cgggccccag	cccctcacct	3240
tctcccctag	ttggggtgga	ccaaagagcc	tgctgttcc	cgccccacct	ggggaagtgg	3300
ggaccacgcc	ttctgctcca	ccccaacgca	atcggaggaa	atctgttcac	cgagtgttgg	3360
cggaaactgga	tgatgagagt	gagcctcctg	agaaccggcc	accggtcctt	atggagocca	3420
ggaagaaact	cogtgtagac	aaagccccac	tgactcccac	tgggtaagtg	gagtcctcac	3480
ttggccctct	cagtgtttta	ctgcttttgc	attccttgta	tccctaggct	gtgaggagggt	3540
ccccctgcct	gggggggatgg	gcacgggagg	tggaatagat	ggaatggcaa	gacctgggtt	3600
agctctgata	gggaaagaaa	aatatgtgca	ggagaacatg	agaggtgggg	tggggcagtg	3660
ctataaaaca	accggagtga	gcatgtcctg	ctttttacat	tcatatggct	ttaaccccac	3720
tttctagtgc	ctaaggatgg	ggaactttca	ggctcacact	agaggttttt	agggccaccc	3780

catgtgtttt taaggacaga gtccaggctc accttagttc tcagaccact gtgcctctgt	3840
ggcctcacc tatgaccagc cataggggtg caaggtctag gccttctcag atttccgggtg	3900
acccttgtgt ctctctcact tccttcagaa atgcacgtgg ccgtcctcgg aagtaccag	3960
tgagcgctcc catggctccc cctgcagttg gggcggggag ccctgtgcag ctcttctgtg	4020
ctgcctgccc caggaagaga cagtggcctg ggttcagtgt gatggctgtg acgtctggtt	4080
ccatgtggcc tgtgttggt gcagcatcca ggctgccagg gaggccgact tccgatgcc	4140
agggtgccg gctggcatc agcctaaggt ccaccgcca ggcaccatcg gacacacctg	4200
cccatgagta gacacagcag cgagcaaata ggtctgataa atacccccct tcccttccct	4260
ccccaaagg gaatgactac aggaagaag gatggattga tgtggactca ttcagggcct	4320
ggagcagacc ctggtggcca agacagaaga gatggtttcc ttccaaagat attgccacct	4380
ccaggaaatt gccagtgagc tggaaattcc cactattaca agccataagg ccatgtcgcc	4440
atggacacca gaatatctgt agtcagagca cctatcagtt gcaaaagcca tgctgcaac	4500
cgatggaaaa tgtaagagg agttcttaag gttcttgggtg gaatcaccca aggtattctg	4560
ggaaaacctt gggcctggcc ccaaaacttc cctactctgt ggctagtect gctgccaaca	4620
aaatcgtagc gacctggctt ttcacagctt tgcttttatt tccaagtcaa ggacaagccg	4680
cttcattcac tcttgggcat ttactcttct tgtgggtctg tgatattcct tgctttccag	4740
ggagaatgtg cttggcaagg tctggagaac taattcagaa tcttagggga aggggagaga	4800
tggaaatata aacctgttta ctggaaagg gcaaaatatg ggttgagctg gaggtaggaa	4860
tacaggtaat taaggtttct agtttaagg aaaacagatc tatttgccat ttaaataagg	4920
taactgggat ttggttaagt tcaccagat agcagaagat ttatttacag gcttcacctg	4980
tactgtcagg gacaagagaa aagcctggta aaccaagcta cagcagttta ccagtgtgat	5040
ggctctcaca cagctccacc ccccggtgg acacagcaga gggcacctgg gctggcctgg	5100
ttcagtgtga atcaaacgc ttaaccaca catggtacat gtgattttct tttgtgagcc	5160
ttacaccaag ccaaactatt gtcaaagcat catttctata gaaataaagc cttatcttga	5220
cctgttctat taaaacctgc cacatccgcc ctttctacc tagatttaat gagcccaagt	5280
ttttttacat ggaagaaatg actctggggc aaagaccct aatgaactag tggcagagcc	5340
aggaataaaa cttgagtaac taatgagtca cttatgggca gagtatgcaa aaaccttaag	5400
tggaaaccaa atagacctg gtatcaagaa agcacaaaagt attaatagaa gtttctggtt	5460
ggggtgatct aggttcaaca gaaataagat gatttctaag tataaagctc aaaattgaat	5520
tc	5522

<210> 25
 <211> 2713
 <212> ADN
 <213> *Mus musculus*
 <400> 25

cagcacggag ggagcgagat ccaggaatct gcaacagaaa ccatgacagc ctgaaacacc	60
ctgtggtgcc aacctccaaa ttctcatctg tcacttcaga ccctgactgg ctgacagagc	120
agcagaatth caactccaat aaacgtgaat gtgcttctag gcaaagcagc caagctgacg	180
agggaggggg gtggaagagc tagctcctct tgggcatttg tcaaactttt acctcctggc	240
tgtgtgcaag gaggggactc aacttcggct tcaagctacc aaggctctgg atccagccac	300
ctctccgcag atctagccag catgaaggct gtgcttctcc tcctgtgcgc cttgggaacc	360
gctgtggcaa tcccgacaag tacaaggtht ctctctgacc actccaaccc aactactgca	420
acactggtga caccggaaga cgctacagtc ccattgccc gggttgaagc tacagcagac	480
atagaaaacc atcccaatga caaggctgaa aaaccttcag cacttaattc agaagaggaa	540
actcatgaac agtcaacaga gcaggacaaa acctacagct tcgaggtgga cctgaaggat	600
gaggaggatg gagatgggga ttttaagtgt gatccaacgg aaggaaact aactctggat	660
ctacaagaag gtacaagtga gcctcaacag aaaagtctcc cggagaacgg cgatttcccc	720
gcgaccgtgt ccacttccta tgtggatcct aaccaacggc caaacatcac aaagggaaag	780
gagagtcagg agcaacctgt aagtactca caccagcaac cgaatgaaag cagcaagcaa	840
acccaagact taaaggctga agaaagccag acacaagatc cagacattcc caatgaagaa	900
gaggaagaag aagaggaaga agaagaggaa gaagaggaag agccggaaga cattggtgcc	960
cccagtata accaagagga gggaaaagaa cctctggagg agcagcctac cagcaagtgg	1020
gaaggaaaca gagagcaatc tgatgacacc ttagaagagt ccagtacgcc cactcagata	1080
agcaagacag agaagcatca atctgagcaa ggaaaccaag ggcaggagag tgactctgag	1140
gcagaaggag aggacaaggc ttcaggcagc aaggaacaca ttccacatac agagcagcag	1200
gaccaagaag ggaaagctgg cctgaagct attggcaacc agaaggacac tgatgagaag	1260
gccgtttcca cagaacctac cgatgctgcc gtggtgccta ggagtcacgg aggagctggt	1320
gataacgggg gcggggatga ctctaagcat ggtgcaggcg atgactactt catccccagc	1380
caggaattcc tagaggccga aaggatgcat tccctctctt attacctcaa atatggcggc	1440
ggcgaggaga caacgactgg cgagagttag aaccagaggg aggctgcaga caaccaagag	1500
gccaaagaaag ctgagagctc accaaatgct gaaccttcag atgagggcaa ctcaaggag	1560
cacagtgtctg gttcttgcaac gaacttccaa tgtaaaaggg gacacatttg caaaaccgat	1620
ccacaaggga aacctcactg tgthtgccaa gatccagaga cttgtcccc tgcaaaaatc	1680
ctagatcagg cttgtggcac tgacaaccaa acctacgcca gctcctgtca cctgtttgct	1740

accaagtgc	ggctggagg	gacaaaaag	ggacaccaac	tgcagctgga	ttacttcgga	1800
gcttgcaaat	ctattcctgc	ttgtacggac	tttgaagtgg	ctcagtttcc	cctgcggatg	1860
agagactggc	tcaaaaacat	cctcatgcag	ctttatgaac	caaataccaa	acatggcggc	1920
tatctcaatg	aaaagcaaag	aagcaaagtc	aaaaaaat	acctggatga	gaagagactc	1980
ttggctggag	accatcccat	tgaacttctc	ttgagggact	ttaagaaaaa	ctaccacatg	2040
tatgtgtatc	ctgtgcactg	gcagtttaat	gaactggacc	agcatcctgc	agacaggatc	2100
ttgacacact	ctgaacttgc	tcctctgcga	gcttcctctg	tgcccatgga	acactgcata	2160
actcgttct	ttgaggagt	tgaccccaac	aaggataagc	acatcacctt	gaaggaatgg	2220
ggccactgct	ttggaattaa	agaggaggat	atagatgaaa	acctcctctt	ttgaattaa	2280
atttgagaga	atcggaactt	tccatccacc	tcacctgctt	taaccgcttc	agaaatacga	2340
gcagccatga	cactatacat	tcatatgtag	caaaacattt	gtttggcatg	tgagagaaga	2400
caatggtagt	aattacttct	tggtgatata	tatatgagcc	aggcacttaa	tattaactta	2460
ggaaatgaaa	ctttaaaatt	aagtagagtc	aatgtctata	aaagactgtc	ctgtctgggg	2520
acagttagcc	accatggcaa	tgtcactctg	tgcatctgcg	tttataattg	ataattataa	2580
actattaaaa	aaacaatgtt	catattgtcc	ataatacctt	atgcatgctg	aggaagtga	2640
atactgtctt	tttgagataa	atatgcctcc	ttttcagtgt	cttgatgtcc	taataaaaaa	2700
tctataaaac	ccc					2713

<210> 26
 <211> 2734
 <212> ADN
 <213> *Mus musculus*

<400> 26

cagcacggag	ggagcgagat	ccaggaatct	gcaacagaaa	ccatgacagc	ctgaaacacc	60
ctgtggtgcc	aacctccaaa	ttctcatctg	tcacttcaga	ccctgactgg	ctgacagagc	120
agcagaattt	caactccaat	aaacgtgaat	gtgcttctag	gcaaagcagc	caagctgacg	180
aggagggggg	gtggaagagc	tagctcctct	tgggcatttg	tcaaactttt	acctcctggc	240
tgtgtgcaag	gagggggactc	aacttcggct	tcaagctacc	aaggctctgg	atccagccac	300
ctctccgcag	atctagccag	catgaaggct	gtgcttctcc	tcctgtgcgc	cttggaacc	360
gctgtggcaa	tcccgacaag	tacaaggttt	ctctctgacc	actccaaccc	aactactgca	420
acactggtga	caccggaaga	cgctacagtc	cccattgccg	gggttgaaagc	tacagcagac	480
atagaaaacc	atcccaatga	caaggctgaa	aaaccttcag	cacttaattc	agaagaggaa	540
actcatgaac	agtcaacaga	gcaggacaaa	acctacagct	tcgaggtgga	cctgaaggat	600
gaggaggatg	gagatgggga	tttaagtgtg	gatccaacgg	aaggaacact	aacactggat	660

ctacaagaag gtacaagtga gcctcaacag aaaagtctcc cggaaaacgg ggattttcccc	720
gcgaccgtgt ccacttccta tgtggatcct aaccaacgcy caaacatcac aaagggaaag	780
gagagtcagg agcaacctgt aagtgactca caccagcaac cgaatgaaag cagcaagcaa	840
acccaagact taaaggctga agaaagccag acacaagatc cagacattcc caatgaagaa	900
gaggaagaag aagaggaaga agaagaggaa gaagaggaag agccggaaga cattggtgcc	960
cccagtata accaagagga gggaaaagaa cctctggagg agcagcctac cagcaagtgg	1020
gaaggaaaca gagagcaatc tgatgacacc ttagaagagt ccagtcagcc cactcagata	1080
agcaagacag agaagcatca atctgagcaa ggaaaccaag ggcaggagag tgactctgag	1140
gcagaaggag aggacaaggc ttcaggcagc aaggaaacaca ttccacatac agagcagcag	1200
gaccaagaag ggaaagctgg ccttgaagct attggcaacc agaaggacac tgatgagaag	1260
gccgtttcca cagaacctac cgatgctgcc gtggtgccta ggagtcacgg aggagctggt	1320
gataacgggg ggggggatga ctctaagcat ggtgcaggcg atgactactt catccccagc	1380
caggaattcc tagaggccga aaggatgcat tccctctcct attacctcaa atatggcggg	1440
ggcgaggaga caacgactgg cgagagttag aaccagaggg aggctgcaga caaccaagag	1500
gccaagaaag ctgagagctc accaaatgct gaaccttcag atgagggcaa ctcaaggag	1560
cacagtgtg gttcttgac gaacttccaa tgtaaaaggg gacacatttg caaaaccgat	1620
ccacaaggga aacctcactg tgtttgccaa gatccagaga cttgtcccc tgcaaaaatc	1680
ctagatcagg cttgtggcac tgacaaccaa acctacgcca gctcctgtca cctgtttgct	1740
accaagtga ggctggaggg gacaaaaaag ggacaccaac tgcagctgga ttacttcgga	1800
gcttgcaaat ctattcctgc ttgtacggac tttgaagtgg ctgagtttcc cctgcggatg	1860
agagactggc tcaaaaacat cctcatgcag ctttatgaac caaatccaa acatggcggc	1920
tatctcaatg aaaagcaaag aagcaaagtc aaaaaaattt acctggatga gaagagactc	1980
ttggctggag accatcccat tgaacttctc ttgagggact ttaagaaaaa ctaccacatg	2040
tatgtgtatc ctgtgcactg gcagtttaat gaactggacc agcatcctgc agacaggatc	2100
ttgacacact ctgaacttgc tcctctgcga gcttccttgg tgcccatgga acactgcata	2160
actcgcttct ttgaggagtg tgacccaac aaggataagc acatcacctt gaaggaatgg	2220
ggccactgct ttggaattaa agaggaggat atagatgaaa acctcctctt ttgaattaa	2280
atttgagaga atcggaactt tccatccacc tcacctgctt taaccgcttc agaaatacga	2340
gcagccatga cactatacat tcatatgtag caaaacattt gtttggcatg tgagagaaga	2400
caatggtagt aattacttct tggatgata tatatgagcc aggcacttaa tattaactta	2460
ggaaatgaaa ctttaaaatt aagtagagtc aatgtctata aaagactgtc ctgtctggg	2520

acagttagcc accatggcaa tgtcactctg tgcactctcg tttataattg ataattataa 2580
actattaaaa aaacaatggt catattgtcc ataatacctt atgcatgctg aggaagtgag 2640
atactgctct tttgagataa atatgcctcc ttttcagtgt cttggatgtc ctaataaaaa 2700
atctataaaa cccccaaaaa aaaaaaaaaa aaaa 2734

<210> 27
<211> 1542
<212> ADN
<213> *Homo sapiens*

<400> 27

cgaccggcac gttcacccca tccctcaggc tttatttatt ttttttcgac aggttctttt 60
caaggctcca gtcaccgcag cagttgtcca tgctgtagtt tccactttcc tgtatgggag 120
ggctggtag gattccactt tcccccaagt gcttagccca gggccagaca aaaagtagtt 180
gcttaagaaa tacttggtga aggaataaat taatgaatga atttggtctt acagcggctg 240
gatggcagac aaacctatga ttataggaac atcaggatct catttggaac agattacgga 300
tgctgcattg tggaacttgg aaaaacaaga gttcttggac aggtttcctg tgaacttgtg 360
tctccaaaac tcaatcgggc aacagaaggt attctttttt taaccttgaa ctctctcaga 420
tggccgctcc agctttcgaa cctggcaggc agtcagatct cttggtgaag ttgaatcgac 480
tcattgaaaag atgtctaaga aattcgaagt gtatagacac tgagtctctc tgtgtgttg 540
ctggtgaaaa ggtttggcaa atacgtgtag acctacattt attaaatcat gatggaaata 600
ttattgatgc tgccagcatt gctgcaatcg tggccttatg tcatttcoga agacctgatg 660
tctctgtcca aggagatgaa gtaacactgt atacacctga agagcgtgat cctgtacat 720
taagtatcca ccacatgccc atttggtgca gttttgcctt tttccagcaa ggaacatatt 780
tattggtgga tcccaatgaa cgagaagaac gtgtgatgga tggcttgctg gtgattgcca 840
tgaacaaaca tcgagagatt tgtactatcc agtccagtgg tgggataatg ctactaaaag 900
atcaagttct gagatgcagt aaaatcgctg gtgtgaaagt agcagaaatt acagagctaa 960
tattgaaagc tttggagaat gacaaaaag taaggaaaga aggtggaaag tttggttttg 1020
cagagtctat agcaaataca aggatcacag catttaaaat ggaaaaggcc cctattgata 1080
cctcggtatg agaagaaaaa gcagaagaaa tcattgctga agcagaacct ccttcagaag 1140
ttgtttctac acctgtgcta tggactcctg gaactgccc aattggagag ggagtagaaa 1200
actcctgggg tgatcttgaa gactctgaga aggaagatga tgaaggcggg ggtgatcaag 1260
ctatcattct tgatggtata aaaatggaca ctggagtaga agtctctgat attggaagcc 1320

aagatgctcc cataatactc tcagatagtg aagaagaaga aatgatcatt ttggaaccag 1380
acaagaatcc aaagaaaata agaacacaga ccaccagtgc aaaacaagaa aaagcaccaa 1440
gtaaaaagcc agtgaaaaga agaaaaaaga agagagctgc caattaaagc taacagttgt 1500
atatctgtat atataactat taaaagggat atttattcca tt 1542

<210> 28
<211> 2808
<212> ADN
<213> *Homo sapiens*

5

<400> 28
cggcatgaga ggccagcctg ccagggaaat ccaggaatct gcaacaaaaa cgatgacagt 60
ctgaaatact ctctggtgcc aacctccaaa ttctcgtctg tcaacttcaga cccccactag 120
ttgacagagc agcagaatat caactccagt agacttgaat gtgcctctgg gcaaagaagc 180
agagctaacg aggaaaggga tttaaagagt ttttcttggg tgtttgtcaa acttttattc 240
cctgtctgtg tgcagagggg attcaacttc aattttctgc agtggctctg ggtccagccc 300
cttacttaaa gatctggaag gcatgaagac tgggcctttt ttcctatgtc tcttgggaac 360
tgcagctgca atcccagaca atgcaagatt attatctgat cattccaaac caactgctga 420
aacggtagca cctgacaaca ctgcaatccc cagtttatgg gctgaagctg aagaaaatga 480
aaaagaaaca gcagtatcca cagaagacga ttcccaccat aaggctgaaa aatcatcagt 540
actaaagtca aaagaggaaa gccatgaaca gtcagcagaa cagggaaga gttctagcca 600
agagctggga ttgaaggatc aagaggacag tgatggtcac ttaagtgtga atttgagta 660
tgcaccaact gaaggtacat tggacataaa agaagatatg attgagcctc aggagaaaaa 720
actctcagag aacactgatt ttttggctcc tgggtgttagt tccttcacag attctaacca 780
acaagaaagt atcacaaga gagaggaaaa ccaagaacaa cctagaaatt attcacatca 840
tcagttgaac aggagcagta aacatagcca aggcctaagg gatcaaggaa accaagagca 900
ggatccaaat atttccaatg gagaagagga agaagaaaaa gagccaggtg aagttggtac 960
ccacaatgat aaccaagaaa gaaagacaga attgcccagg gagcatgcta acagcaagca 1020
ggaggaagac aatacccaat ctgatgatat tttggaagag tctgatcaac caactcaagt 1080
aagcaagatg caggaggatg aatttgatca gggtaaccaa gaacaagaag ataactcaa 1140
tgcagaaatg gaagaggaaa atgcatcgaa cgtcaataag cacattcaag aaactgaatg 1200
gcagagtcaa gagggtaaaa ctggcctaga agctatcagc aaccacaaag agacagaaga 1260
aaagactgtt tctgaggctc tgctcatgga acctactgat gatggtaata ccacgcccag 1320
aaatcatgga gttgatgatg atggcgatga tgatggcgat gatggcgga ctgatggccc 1380
caggcacagt gcaagtgatg actacttcat cccaagccag gcctttctgg aggccgagag 1440

```

agctcaatcc attgcctatc acctcaaaat tgaggagcaa agagaaaaag tacatgaaaa 1500
tgaaaatata ggtaccactg agcctggaga gcaccaagag gccagaaaag cagagaactc 1560
atcaaagtag gaggaacgt caagtgaagg caacatgagg gtgcatgctg tggattcttg 1620
catgagcttc cagtgtaaaa gaggccacat ctgtaaggca gaccaacagg gaaaacctca 1680
ctgtgtctgc caggatccag tgacttctcc tccaacaaaa ccccttgatc aagtttgtgg 1740
cactgacaat cagacctatg ctagttcctg tcatctatc gctactaaat gcagactgga 1800
ggggacaaaa aaggggcatc aactccagct ggattatctt ggagcctgca aatctattcc 1860
tacttgtacg gactttgaag tgattcagtt tcctctacgg atgagagact ggctcaagaa 1920
tatcctcatg cagctttatg aagccaactc tgaacatgct ggttatctaa atgagaagca 1980
gagaaataaa gtcaagaaaa ttacctgga tgaaaagagg cttttggctg gggaccatcc 2040
cattgatctt ctcttaaggg actttaagaa aaactaccac atgtatgtgt atcctgtgca 2100
ctggcagttt agtgaacttg accaacaccc tatggataga gtcttgacac attctgaact 2160
tgctcctctg cgagcatctc tgggtgccat ggaacactgc ataaccggtt tctttgagga 2220
gtgtgacccc aacaaggata agcacatcac cctgaaggag tggggccact gctttggaat 2280
taaagaagag gacatagatg aaaatctctt gttttgaacg aagattttaa agaactcaac 2340
tttcagcat cctcctctgt tctaaccact tcagaaatat atgcagctgt gatacttgta 2400
gatttatatt tagcaaatg ttagcatgta tgacaagaca atgagagtaa ttgcttgaca 2460
acaacctatg caccaggtat ttaacattaa ctttggaac aaaaatgtac aattaagtaa 2520
agtcaacata tgcaaaatac tgtacattgt gaacagaagt ttaattcata gtaatttcac 2580
tctctgcatt gacttatgag ataattaatg attaaactat taatgataaa aataatgcat 2640
ttgtattgtt cataatatca tgtgcacttc aagaaaatgg aatgctactc ttttgtggtt 2700
tacgtgtatt attttcaata tcttaatacc ctaataaaga gtccataaaa atccaaaaaa 2760
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 2808

```

<210> 29
 <211> 2645
 <212> ADN
 <213> *Homo sapiens*

<400> 29

```

gagcagcaga atttcaactc cagtagactt gaatatgcct ctgggcaaag aagcagagct 60
aacgaggaaa gggatttaaa gagtttttct tgggtgtttg tcaaactttt attccctgtc 120
tgtgtgcaga ggggattcaa cttcaatttt tctgcagtgg ctctgagtcc agccccttac 180
ttaaagatct ggaaagcatg aagactgggc ttttttctct atgtctcttg ggaactgcag 240

```

ctgcaatccc gacaaatgca agattattat ctgatcattc caaaccaact gctgaaacgg	300
tagcaccoga caacactgca atccccagtt taagggctga agatgaagaa aatgaaaaag	360
aaacagcagt atccacagaa gacgattccc accataaggc tgaaaaatca tcagtactaa	420
agtcaaaaga ggaaagccat gaacagtcag cagaacaggc caagagttct agccaagagc	480
tgggattgaa ggatcaagag gacagtgatg gtgacttaag tgtgaatttg gagtatgcac	540
catctgaagg tacattggac ataaaagaag atatgagtga gcctcaggag aaaaaactct	600
cagagaacac tgattttttg gctcctgggt ttagttcctt cacagattct aaccaacaag	660
aaagtatcac aaagagagag gaaaaccaag aacaacctag aaattattca catcatcagt	720
tgaacaggag cagtaaacat agccaaggcc taagggatca aggaaacca gagcaggatc	780
caaatatctc caatggagaa gaggaagaag aaaaagagcc aggtgaagtt ggtaccocaca	840
atgataacca agaaagaaag acagaattgc ccaggagca tgctaacagc aagcaggagg	900
aagacaatac ccaatctgat gatatttttg aagagtctga tcaaccaact caagtaagca	960
agatgcagga ggatgaattt gatcagggtg accaagaaca agaagataac tccaatgcag	1020
aatggaaga ggaaaatgca tcgaacgtca ataagcacat tcaagaaact gaatggcaga	1080
gtcaagaggg taaaactggc ctagaagcta tcagcaacca caaagagaca gaagaaaaga	1140
ctgtttctga ggctctgctc atggaaccta ctgatgatgg taataccacg ccagaaatc	1200
atggagttga tgatgatggc gatgatgatg gcgatgatgg cggcactgat ggccccaggc	1260
acagtgcaag tgatgactac ttcatcccaa gccaggcctt tctggaggcc gagagagctc	1320
aatccattgc ctatcacctc aaaattgagg agcaaagaga aaaagtacat gaaaatgaaa	1380
atataggtac cactgagcct ggagagcacc aagaggccaa gaaagcagag aactcatcaa	1440
atgaggagga aacgtcaagt gaaggcaaca tgagggtgca tgctgtggat tcttgcata	1500
gcttccagtg taaaagaggc cacatctgta aggcagacca acagggaata cctcactgtg	1560
tctgccagga tccagtgact tgtcctccaa caaaaccctt tgatcaagtt tgtggcactg	1620
acaatcagac ctatgctagt tcctgtcatc tattcgctac taaatgcaga ctggagggga	1680
ccaaaaaggg gcatcaactc cagctggatt attttgagc ctgcaaatct attcctactt	1740
gtacggactt tgaagtgatt cagtttcctc tacggatgag agactggctc aagaatatcc	1800
tcatgcagct ttatgaagcc aactctgaac acgctggtta tctaaatgag aagcagagaa	1860
ataaagtcaa gaaaatttac ctggatgaaa agaggctttt ggctggggac catcccattg	1920
accttctctt aagggacttt aagaaaaact accacatgta tgtgtatcct gtgcactggc	1980
agtttagtga acttgaccaa caccctatgg atagagctct gacacattct gaacttgctc	2040
ctctgcgagc atctctggtg cccatggaac actgcataac ccgtttcttt gaggagtgtg	2100
acccaacaa ggataagcac atcacctga aggagtgggg ccactgcttt ggaattaaag	2160

aagaggacat agatgaaaat ctcttgTTTT gaacgaagat tttaaagaac tcaactttcc	2220
agcatcctcc tctgttctaa ccacttcaga aatatatgca gctgtgatac ttgtagattt	2280
atatttagca aaatgttagc atgtatgaca agacaatgag agtaattgct tgacaacaac	2340
ctatgcacca ggtattttaac attaaccttg gaaacaaaaa tgtacaatta agtaaagtca	2400
acatatgcaa aatactgtac attgtgaaca gaagtttaat tcatagtaat ttcactctct	2460
gcattgactt atgagataat taatgattaa actattaatg ataaaaataa tgcatttgta	2520
ttgttcataa tatcatgtgc acttcaagaa aatggaatgc tactcttttg tggtttacgt	2580
gtattatttt caatatctta atacccta atacccta atacccta taaaaatcca aaaaaaaaaa	2640
aaaaa	2645

<210> 30
 <211> 2808
 <212> ADN
 <213> *Homo sapiens*

<400> 30

cggcatgaga ggccagcctg ccagggaaat ccaggaatct gcaacaaaaa cgatgacagt	60
ctgaaatact ctctggtgcc aacctccaaa ttctcgtctg tcaacttcaga cccccactag	120
ttgacagagc agcagaatat caactccagt agacttgaat gtgcctctgg gcaaagaagc	180
agagctaacg aggaaaggga tttaaagagt ttttcttggg tgtttgtcaa acttttattc	240
cctgtctgtg tgcagagggg attcaacttc aattttctgc agtggctctg ggtccagccc	300
cttacttaaa gatctggaaa gcatgaagac tgggcctttt ttccatgtc tcttgggaac	360
tgcagctgca atcccgacaa atgcaagatt attatctgat cattccaaac caactgctga	420
aacggtagca cctgacaaca ctgcaatccc cagtttatgg gctgaagctg aagaaaatga	480
aaaagaaaca gcagtatcca cagaagacga ttcccaccat aaggctgaaa aatcatcagt	540
actaaagtca aaagaggaaa gccatgaaca gtcagcagaa cagggaaga gttctagcca	600
agagctggga ttgaaggatc aagaggacag tgatggtcac ttaagtgtga atttgagta	660
tgcaccaact gaaggtacat tggacataaa agaagatatg attgagcctc aggagaaaaa	720
actctcagag aacactgatt ttttggtccc tgggtgttagt tccttcacag attctaacca	780
acaagaaagt atcacaaaga gagaggaaaa ccaagaacaa cctagaaatt attcacatca	840
tcagttgaac aggagcagta aacatagcca aggcctaagg gatcaaggaa accaagagca	900
ggatccaaat atttccaatg gagaagagga agaagaaaaa gagccaggtg aagttggtac	960
ccacaatgat aaccaagaaa gaaagacaga attgcccagg gagcatgcta acagcaagca	1020
ggaggaagac aatacccaat ctgatgatat tttggaagag totgatcaac caactcaagt	1080


```

aagcaagatg caggaggatg aatttgatca gggtaaccaa gaacaagaag ataactccaa 1140
tgcagaaatg gaagaggaaa atgcatcgaa cgtcaataag cacattcaag aaactgaatg 1200
gcagagtcaa gagggtaaaa ctggcctaga agctatcagc aaccacaaaag agacagaaga 1260
aaagactgtt tctgaggctc tgctcatgga acctactgat gatggtaata ccacgccag 1320
aatcatgga gttgatgatg atggcgatga tgatggcgat gatggcggca ctgatggccc 1380
caggcacagt gcaagtgatg actacttcat cccaagccag gcctttctgg aggccgagag 1440
agctcaatcc attgcctatc acctcaaaat tgaggagcaa agagaaaaag tacatgaaaa 1500
tgaaaaatata ggtaccactg agcctggaga gcaccaagag gccaagaaag cagagaactc 1560
atcaaatgag gaggaacgt caagtgaagg caacatgagg gtgcatgctg tggattcttg 1620
catgagcttc cagtgtaaaa gaggccacat ctgtaaggca gaccaacagg gaaaacctca 1680
ctgtgtctgc caggatccag tgacttgtcc tccaacaaaa ccccttgatc aagtttgtgg 1740
cactgacaat cagacctatg ctagttctctg tcatctattc gctactaaat gcagactgga 1800
ggggacaaa aaggggcac cactccagct ggattatctt ggagcctgca aatctattcc 1860
tacttgtacg gactttgaag tgattcagtt tcctctacgg atgagagact ggctcaagaa 1920
tacctcatg cagctttatg aagccaactc tgaacatgct ggttatctaa atgagaagca 1980
gagaaataaa gtcaagaaaa tttacctgga tgaaaaggag cttttggctg gggaccatcc 2040
cattgatctt ctcttaaggg actttaagaa aaactaccac atgtatgtgt atcctgtgca 2100
ctggcagttt agtgaacttg accaacaccc tatggataga gtcttgacac attctgaact 2160
tgctcctctg cgagcatctc tggtgcccat ggaacactgc ataaccggtt tctttgagga 2220
gtgtgacccc aacaaggata agcacatcac cctgaaggag tggggccact gctttggaat 2280
taaagaagag gacatagatg aaaatctctt gttttgaacg aagattttta agaactcaac 2340
tttccagcat cctcctctgt tctaaccact tcagaaatat atgcagctgt gatacttgta 2400
gatttatatt tagcaaatg ttagcatgta tgacaagaca atgagagtaa ttgcttgaca 2460
acaacctatg caccaggtat ttaacattaa ctttggaac aaaaatgtac aattaagtaa 2520
agtcaacata tgcaaaatac tgtacattgt gaacagaagt ttaattcata gtaatttcac 2580
tctctgcatt gacttatgag ataattaatg attaaactat taatgataaa aataatgcat 2640
ttgtattgtt cataatatca tgtgcacttc aagaaaatgg aatgctactc ttttgtggtt 2700
tacgtgtatt attttcaata tcttaatacc ctaataaaga gtccataaaa atccaaaaaa 2760
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2808

```

<210> 31
 <211> 2678
 <212> ADN
 <213> *Coturnix coturnix*
 <400> 31

gtcagacctt cctcctcaga agctcacaga aaaacacgct ttctgaaaga ttccacactc	60
aatgccaaaa tataccacag gaaaattttg caaggctcac ggatttccag tgcaccactg	120
gctaaccaag taggagcacc tcttctactg ccatgaaagg aaaccttcaa accctaccac	180
tgagccatta actaccatcc tgtttaagat ctgaaaaaca tgaagactgt attgctcctg	240
atttgtcttc taggatctgc tttcaccact ccaaccgatc cattgaacta ccaatttggg	300
gcccattggac agaaaactgc agagaagcat aaatatactc attctgaaat gccagaggaa	360
gagaacacag ggtttgtaaa caaaggatgat gtgctgtctg gccacaggac cataaaagca	420
gaggtaccgg tactggatac acagaaggat gagccctggg cttccagaag acaaggacaa	480
ggtgatggtg agcatcaaac aaaaaacagc ctgaggagca ttaacttcct tactctgcac	540
agtaatccag ggttggttc tgataaccag gaaagcaact ctggcagcag cagggaacag	600
cacagctctg agcaccacca gcccaggagg cacaggaaac acggcaacat ggctggccag	660
tgggctctga gaggagaaag tccagtggat gctcttggtc tggttcgtga gcgcaacaca	720
tggaaataca ataaaaacac agttggccta gatgaaaaca acaatggaag tgaagaagag	780
gaagctgggg aggaagaaga tgaggaatgg ggtgaagaaa ctgattacag ggatatgaaa	840
cacagagccc gtgggacaag ccatggaaga gaatacagaa gatggcaaaa tgaaaacagc	900
cggccatctg gtgaattctt gagagattcc agtctgccag tacgtataac caagagacac	960
ggtgagaaat tcagcatgga ggaggaaagt caggaaaagc tctacaagga aggaaaactc	1020
cctctctcaa agaaaaatca taatgaggat caaggtgaaa aaagacaaag tgaagaagt	1080
aaagagcatt ttcaagtagt caatcagcgc aaacacagag cagtgcgaa aaggcaggat	1140
aaggagggca gcaatgctga ggaggatgat aatgatagtg gtgatgatgg tgaggaagat	1200
cttggcaatg tctggaggga agcagtctac gaggaaggag aaagaatgca aagcaatgac	1260
caggacagta tctaatacaa gcaaaaagag gaaataactg ctggagatga cagtggagtt	1320
tatagggaga tgcaggatta caaaggatgac aaaattaaag atgttactca ctctgaagac	1380
aatcattacc accatgagcc ccctaattcc agcagcaagc aacaactgca aacaagtacg	1440
tctgttgaga gcatgaattc aacagagcat gaggatgagg ttaagaccac aggaggttca	1500
tatcatgagg aaagtgcaag gaacagcact gggaaggctc tcccggatct ttgtagaaac	1560
ttccactgca aaagaggaaa agtctgccaa gcagacaagc aaggaaaacc cagctgtatt	1620
tgccaagatc ctgctgcttg cccttccacc aaagattata agcgtgtttg tggcactgat	1680
aataagactt acgatggtac gtgccaaactc tttggcacca aatgtcaact tgaagggaca	1740
aaaatgggac gccagctgca cctggactat atgggtgcct gcaaacacat accccactgt	1800

actgattacg aagtgaatca gttccctctc cgtatgagag actgggtcaa aaacatccta 1860
atgcaatatt atgaacgtga tcaggatacg tctgcatttc taaccgaaaa gcaaaggaat 1920
aagggtcaaaa agatatacct gaatgagaag cgtctcgtct ctggtgagca cccagttgag 1980
cttctcctgc atgactttga gaaaaactac cacatgtatc tctatcctgt gcactggcaa 2040
ttttatcagc ttgaccagca cccagttgac agatcactga ctcatcaga gctcgtcct 2100
ttgagagcct ccctcgttcc catggaacac tgcataaccc gtttcttcca ggagtgtgat 2160
ggagaccaag acaaacttat cactttgaaa gagtgggtgcc actgctttgc gattaaggaa 2220
gaagacataa atgaaaatct cctgttctga gccacactga gcagaatccc catgcagcgc 2280
tacagcttgt caaacatgca atgcccattt atgactgcaa ttaacagctc tgtaatttt 2340
caggaataag ttggcataag attcttgagg gcagaacaag tcgctcttgg ataacacaag 2400
tgcctaattg ttacaaattc attaacagca gtagtgttta agagctctaa gtagctcata 2460
cttaaagagt gtttccctct gcacgtacca ataatctctt agtaagacga ctaacttgat 2520
gactgagttg ttcacaaaac ccttccgtag aattatagga tgtggatttt ataatacacc 2580
gataaaaact actttgaaat aggttttctt ttctgtcgt ttactgtcag tagctctctg 2640
catagaaatg tcaaataaac agatcttgtt ttggtttc 2678

<210> 32
<211> 2823
<212> ADN
<213> *Mus musculus*

<400> 32

aagcgacgct cccaccttcg cctctaactc gctgcgcgca ccctgccag tgtcctcgg 60
agtcccggac ccgagcacga tgtggaaacg atggctggcg ctctcgctgg tgaccatgcg 120
cctgggtccac ggcgaggagg aacctagaag caaatccaag atctgcgcca atgtgttttg 180
tggagctggc agggaatgtg ccgtcacaga gaagggggag cccacgtgcc tctgcattga 240
gcaatgcaaa cctcacaaga ggcctgtgtg tggcagtaat ggcaagacct acctcaacca 300
ctgtgaactt catagagatg cctgcctcac tggatccaag atccaggttg attatgatgg 360
gcactgcaaa gaaaagaagt ctgcgagtc atctgccagc ccagttgtct gctatcaagc 420
taaccgcgat gagctccgac ggcgcctcat ccagtggctg gaagctgaga tcattccaga 480
tggctggttc tctaaaggca gtaactacag tgagatccta gacaagtact ttaagagctt 540
tgataatggc gactctcacc tggactccag tgaattcctg aaattcgtgg agcagaatga 600
aacagccatc aacatcacca cttatgcaga tcaggagaac aacaaactgc tcagaagcct 660
ctgtgttgac gccctcattg aactgtctga tgagaacgct gactggaaac tcagcttcca 720
agagttcctc aagtgcctca acccatcctt caaccctcct gagaagaagt gtgccctgga 780

ggtcgaaacc tatgcagatg gagctgagac tgaggtggac tgcaatcgct gtgtctgttc	840
ctgtggccac tgggtctgca cagcaatgac ctgtgatgga aagaatcaga aggggggtcca	900
gaccacaca gaggaggaga agacaggata tgtccaggaa ctccagaagc accagggcac	960
agcagaaaag accaagaagg tgaacaccaa agagatctaa gaagaggcac agagcaccgt	1020
gtccggagcc cagcgctcc tcttcagcgc tgagcccagt acacacagag tctgcagcaa	1080
tcaccaaadc actagtattt gcttgatgg cagcgaatct tattttgttt gttttgcaat	1140
aaaggaaatg aggggtggcca gcctagcgag ggaaggccac aaccttcacc tgtaggaatg	1200
ctttaagaga aactaaagga caccttggga cgagaggcaa ctaaggaaac agcatcgggt	1260
tggcagagga gcagaggcag gtttgatga agcctttctg gggtcacagc agctgcgag	1320
agaatacagg aaaagcatag agaaacattg aactagccct gctggaggaa gtggggggag	1380
ctttgtaggg aggaaccctg ctgctttgac ccttgcacc actgtcagca tgacagacct	1440
gcagcaagtc tgcttctcct tttggccca acaatcacct gaacacacag ccgcccact	1500
agttacctgt gtcctcagcc ttgcatggag tttcctggag gaggtgttta aatgatgcag	1560
acacttatgt acttcaagcg catggagact aaccaaattt ttaaataca tttttctttt	1620
tttttttttt ttgttaacca aaggtgctat ttctctgtaa aagacttttt tccaagctga	1680
cttcattcct cagttattac cgttatatta ttgtgtttt ttaatatctt attttttgac	1740
tagatattaa gcttttgtaa ttatttttca ttagtcttac tatttcgaga agtgaagggtg	1800
aaggggggtt gggcattttt ccagggtaca ggggaactctg taacacaaac agcccatacc	1860
ctgtcacata ttagaccggt tgcagttcgg agcatgcacc ccaaccaga gcttctagaa	1920
aatcagctcc atgccacgaa ggcacaagag gccctcagc agaagccaca ggacaaagca	1980
tcttcataga cagctgtga gatccaaaca gtttaatttg tttgtttct tgtaagaagt	2040
tccaaggatg gacgctcagg ctatcccagc ctgccagcct gctgtgatct gtggctaact	2100
ggcagagtca gccactgtgg tccttagctg ctctgtttc taggtgtcag tttacttagt	2160
aaactggtaa gaatgaatct tggaatttaa taaatggtag tttgtggtt agccaactgg	2220
tccagaggga gctaccttct ccttaggata gatgaatcta ctccataaga aaaaccagcc	2280
aggaatagca tggatgggtt ttgctttggt tgaaatgatc ctagcagggtg actgggtatg	2340
aggacttcac ggtcactctg ccaggaaga gagcgtgaag gacaactagc agcttcctta	2400
gggatggtac acatgtgtgt gatctctgga gatcagaggt tgccccacac acatgatgat	2460
aaaacttttc agatttagag cggttaaaac tggagatcga atctggattg agaatcagca	2520
ctgggggcag aaactgttat tgaaagtcaa tcctttcttt gagacactcc gaataaacta	2580
tggagatttt cctgcatagg aaagtgtgga atgttgagct attgagatgg gagtggaatt	2640

cgtoctaaat agttttttcc tggctctcatc tgaacaagac aatttgctct gcctagtgtt	2700
ctgtgccctc cttttcaaaa gctctgagcc ccgctcatgc agtccagatt tcatccccct	2760
ctccaagtgc cttggagagc tcacgacagc aatgccatca tcaaaagttt tgctgctggg	2820
aag	2823

<210> 33
 <211> 3484
 <212> ADN
 <213> *Homo sapiens*

5

<400> 33

cactctctgt tgtccaatgg acacacctgt cgtgttttga gccagcgaga gatgcagtgg	60
aagtgaaaag catggttaca gactcccat gcgacagtac actcttctga agtagcggac	120
gcctgggttag cttgacattc tatgcaaaga tccataatgt ggttcctgca gatggcacag	180
ttatcaacca caatatccca ggcccgagag gctactgcat tccacttttt cacttcaaag	240
cgcttcttgc ccgcgccgct gttggtgccg ctcggggtat ccacatccat cgctgcgggc	300
tcacaaagcg gccagacgct cggcgggcggc gtgtggcagg agcgcagggg cgcgagccgg	360
cgatcagcct tcccgccgac cgtgcccgcg gagctcgagc aactcggact aggggacccg	420
ggcgggcccc caagatgccg gcgacgcggc tgttggcggc ggccgcgcgc gcgtgggtgct	480
tcctccaagt cgagagccgg cacctggacg cgctcgccgg aggcgcgggc cccaaccacg	540
gcaatttcct agacaatgac cagtggctga gcaccgtctc ccagtacgac cgggacaagt	600
actggaaccg ctttcgagac gatgattatt tcagaaactg gaatcccaac aagccctttg	660
accaagccct ggacccatcc aaggaccctt gcctgaaggt aaaatgcagc cctcaciaag	720
tgtgtgtgac ccaggactac cagaccgccc tgtgtgtcag ccgcaagcac ctgctcccc	780
ggcaaaagaa ggggaacgtg gcccagaaac actgggttgg accttogaat ttggtcaagt	840
gcaagccctg tcccggtgga cagtcagcca tggctctgcg ctcagatggc cactcctaca	900
catccaagtg caaattggag ttccatgctt gttctactgg caaaagcctc gccaccctct	960
gtgatgggccc ctgtccctgt ctcccagagc ctgagccacc aaagcacaag gcagaaagga	1020
gtgcctgcac agacaaggag ttgcggaacc ttgcctcccg gctgaaggat tggtttgag	1080
ctctccacga ggatgcgaac agagtcacac agcccaccag ctccaacaca gcccaaggca	1140
ggtttgacac tagcatcctg cccatctgca aggactccct gggttggtat ttcaacaagt	1200
tggacatgaa ctatgacctc ctgcttgacc cttcagagat caatgccatc tacctggata	1260
agtacgagcc ctgtatcaag cctcttttca actcgtgtga ctcttcaag gatggcaagc	1320

10

tttctaacaa tgagtggcgc tactgcttcc agaagcctgg aggtctccct tgccagaatg	1380
aatgaacag aattcagaag ctgagtaagg ggaaaagcct gttggggggcc ttcatacctc	1440
ggtgtaatga ggagggtat taaaaagcca cacagtgccg cggcagcacg gggcagtgc	1500
ggtgtgtgga caaatatggg aatgagttgg ctggctccag gaaacagggg gctgtgagct	1560
gtgaagagga gcaggaaacc tcaggggatt ttggcagtgg tgggtccgtg gtcctgctgg	1620
atgacctaga atatgaacgg gagctgggac caaaggacaa agaggggaag ctgagggcgc	1680
acacccgagc cgtgacagag gatgatgagg atgaggatga tgacaaagag gatgaggtcg	1740
ggtacatatg gtagtgccca caagaaagag gacacaagtt ttgcacaaaa ttgcaagtca	1800
cttcctattc ctgcatttgt atctaagact ccaaggcacc aaggtctctt ctccatttgt	1860
gctctctata cccgacctaa ggtttggaag acaactgctt gttcccagag gattctgatt	1920
ttgcatatgt ttgtatggga gaaaggggtg tgtgtttttt tttttgttgt tgtttatttt	1980
ttggataggg aagtcattgg cttaattaga gcctccttcc tttctgtgag atttttccaa	2040
caagcatgtg atttacgtgg aattctgaca gtgcaggag cccccacct cttaaatgtc	2100
aaagaccctt tttgattacc cactctgtgt gttattacag catggttccc agccttacag	2160
tgtctaagtg cttctcttgt gtcctgtaga tgttgtgaaa aagaaaaaaa caaaaaatac	2220
accacactgt actttttccc cctgcccccg ttactgccgg tgattattat taaaaattag	2280
tttttttcac atcattctat ctggcttccct ataaacaaca gccttaattc agtcaagact	2340
ccctttggga attcatttta ttaaaaattg gtgtctggat acttccctgt acatgcataa	2400
atatgcatgc atgtacagaa agactgtatg tgtgtgcctt gcacacacac ccatacctct	2460
cagaaaaagt gtttgggtat cttaaaaact cgaaaaacaa tgataaattt ctcagcttgt	2520
ccagacctgg aacaaaattt ctggaataag aaatttgtat taaagtcctt ttttgacta	2580
acagtgggtc cttgtagcct gcaggctgag gaagtctctt ctctgtgcat cagcagagtt	2640
actgaaagcc tctgattgag aaaaaacctc cgtctgccta aatcaacttt ctcgcagaag	2700
ccatgcgact cccacacgac acgggcagct tcacaagcca tctctttcat ttctgottga	2760
agccccctgg ctgcagcaat cctgtctgcc ataggtttct tccttcctta cctactcaag	2820
ggctttttct aaggcatgca cacatatctc ctgttctctg agagtaccat ggtgttcctt	2880
aaaagaagaa aatttctaatt tctgaactca atgttttgct tttactccct ttctactgac	2940
aatcatgat aagggcacaa aagctgtaca gatttttttt tttaaccact caatcccaaa	3000
tggaggccta caaagaacat cgtaataaca catggaagca aaccccggtt ttttaagagc	3060
aaattctgtc cccccctcac tcccccaagt gacaagatac taatgaagaa agttcttcac	3120
catagtgttt gttttaacta aactcattgg agtctagtcc caaatttggg agggcatca	3180
tctctacatt ccttaggatt tctctccta tcaagctggc ccagatacaa gtaccaaaca	3240

gtagtctctg aagttcccat ttccttcagt accagtttat aagctactgt ccgccactga 3300
 ttttcatcta tcaggggtgtc ctaatcagaa tcagccaccc aagcaagcct ctctggccca 3360
 catatctatc tcttgccctc ccccatgaac ttcagcctgt ccacacaaaa gccacataaa 3420
 ctcaagcaag aaatatgttc agccaaaaca tgattatagt ggcagctgac caatacccca 3480
 cccc 3484

<210> 34
 <211> 1611
 <212> ADN
 <213> *Mus musculus*

5

<400> 34

ggaatccctt cctcagttta ccccgcgagc gaaaaaaga aaggattgtc ggaggtccc 60
 ggatgttgag ccgctgagg atccggcttc actgaggtgc cacaagacgt ctccctggtt 120
 ttgcatogca tcaggacttt actggaattg ctttttaaac ctcatcttga cagcctctgt 180
 cttggtgcat atccaggcct ctggcgctct gaatctgcac tgcttaagtt ggcattaggc 240
 acaatactta gctgttgaca tcagagagaa ctgaattgaa gccacgggtg catgaccag 300
 gagtaggaga ggaatggac ttctgaggtg ggcgaagtgg tgcaagggg gcacccccca 360
 tgctgccctg tttccagctg ctgcgtatag gggcgggcag gggcggtgac ctctacacct 420
 tccaccccc gtccaagtct ggctgcacct atggattggg ctgcagggcc gacctgtgtg 480
 atgtggccct gcggccccag caggagcccg gcctcatctc tggagtccat gcggaattgc 540
 atgtgaact ccaaggggac gactggaggg tcagcctgga ggatcacagc agccaaggga 600
 ctttggtaaa taatgtccga ctccaagag gccacagact ggagttgagt gatggtgacc 660
 ttctgacctt tggccccaa ggtcaagcag gaaccagctc ctctcagag ttctacttca 720
 tgtttcaaca agtccgggtt aaacctcagg actttgcagc cataaccgtc cctcggtcta 780
 agggagaagc tggggggcgt ttccaacctt tgctgcccc ccaaggggca cctcagaggc 840
 cactcagcac tctctcctct gcccccaagg ccacactgat tctcaattcc atcggcagcc 900
 tcagcaaact ccaggcccag cctctcacct tctcccggtg tggtaggcagg ccacagggcc 960
 tggctattcc ctctcagcat ggggaagcgc aagtttcgcc tgctccaccc acaagaaacc 1020
 ggaggaaatc agctcataaa gtgttggcag agctagatga cgaggtctcc ccaggccccc 1080
 tgtecgctct gacggagccc aggaagaggc tccgggtgga gaaagctgct ctgatagcca 1140
 gtggggaatg accacaggga aagggtccatt gctaaagact gactcaaggc tggagaatgc 1200
 ttgcaaata agtgacagaa aagatgggtt cctgccaaag atgttgagc ctccaagttt 1260
 ccagtgaagt acaagcccag tcacaaggaa taaagccttg cctccatgga tgccaggata 1320

10

cctgcagtta gagcaccggt cccagaaacc attcctagaa ggtatggttt tattttgctg 1380
 aagccagcat ataccaggca agtgctctgc cacaataccc taacctact catgttcttg 1440
 aatggcatcg gctaaagcac cctgaaggac ccagagcaca ttccaaggc atcctcatgc 1500
 caccgctagc tagtatcacc accagctgag tctctgggac tagcttttca gagcttagtc 1560
 tttattccca aataaagaat aaactgtttc atccaaaaaa aaaaaaaaaa a 1611

<210> 35
 <211> 265
 <212> PRT
 <213> *Homo sapiens*

<400> 35

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

<210> 36
 <211> 676

<212> PRT

<213> *Coturnix coturnix*

<400> 36

5

```

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

Thr Pro Thr Asp Pro Leu Asn Tyr Gln Phe Gly Ala His Gly Gln Lys
      20      25      30

Thr Ala Glu Lys His Lys Tyr Thr His Ser Glu Met Pro Glu Glu Glu
      35      40      45

Asn Thr Gly Phe Val Asn Lys Gly Asp Val Leu Ser Gly His Arg Thr
50      55      60

Ile Lys Ala Glu Val Pro Val Leu Asp Thr Gln Lys Asp Glu Pro Trp
65      70      75      80

Ala Ser Arg Arg Gln Gly Gln Gly Asp Gly Glu His Gln Thr Lys Asn
      85      90      95

Ser Leu Arg Ser Ile Asn Phe Leu Thr Leu His Ser Asn Pro Gly Leu
      100      105      110

Ala Ser Asp Asn Gln Glu Ser Asn Ser Gly Ser Ser Arg Glu Gln His
      115      120      125

Ser Ser Glu His His Gln Pro Arg Arg His Arg Lys His Gly Asn Met
      130      135      140

Ala Gly Gln Trp Ala Leu Arg Gly Glu Ser Pro Val Asp Ala Leu Gly
145      150      155      160

Leu Val Arg Glu Arg Asn Thr Trp Lys Tyr Asn Lys Asn Thr Val Gly
      165      170      175

Leu Asp Glu Asn Asn Asn Gly Ser Glu Glu Glu Glu Ala Gly Glu Glu
      180      185      190

Glu Asp Glu Glu Trp Gly Glu Glu Thr Asp Tyr Arg Asp Met Lys His
      195      200      205

Arg Ala Arg Gly Thr Ser His Gly Arg Glu Tyr Arg Arg Trp Gln Asn
210      215      220

Glu Asn Ser Arg Pro Ser Gly Glu Phe Leu Arg Asp Ser Ser Leu Pro
225      230      235      240

```

Val	Arg	Ile	Thr	Lys	Arg	His	Gly	Glu	Lys	Phe	Ser	Met	Glu	Glu	Glu			
				245					250					255				
Ser	Gln	Glu	Lys	Leu	Tyr	Lys	Glu	Gly	Lys	Leu	Pro	Leu	Ser	Lys	Lys			
			260					265					270					
Asn	His	Asn	Glu	Asp	Gln	Gly	Glu	Lys	Arg	Gln	Ser	Glu	Glu	Ser	Lys			
		275					280					285						
Glu	His	Phe	Gln	Val	Val	Asn	Gln	Arg	Lys	His	Arg	Ala	Val	Thr	Lys			
	290					295					300							
Arg	Gln	Asp	Lys	Glu	Gly	Ser	Asn	Ala	Glu	Glu	Asp	Asp	Asn	Asp	Ser			
305					310					315					320			
Gly	Asp	Asp	Gly	Glu	Glu	Asp	Leu	Gly	Asn	Val	Trp	Arg	Glu	Ala	Val			
				325					330					335				
Tyr	Glu	Glu	Glu	Glu	Arg	Met	Gln	Ser	Asn	Asp	Gln	Asp	Ser	Ile	Thr			
			340					345					350					
Asn	Lys	Gln	Lys	Glu	Glu	Ile	Thr	Ala	Gly	Asp	Asp	Ser	Gly	Val	Tyr			
		355					360					365						
Arg	Glu	Met	Gln	Asp	Tyr	Lys	Gly	Asp	Lys	Ile	Lys	Asp	Val	Thr	His			
	370					375					380							
Ser	Glu	Asp	Asn	His	Tyr	His	His	Glu	Pro	Pro	Asn	Ser	Ser	Ser	Lys			
385					390					395					400			
Gln	Gln	Leu	Gln	Thr	Ser	Ser	Ser	Val	Glu	Ser	Met	Asn	Ser	Thr	Glu			
				405					410					415				
His	Glu	Asp	Glu	Val	Lys	Thr	Thr	Gly	Gly	Ser	Tyr	His	Glu	Glu	Ser			
			420					425					430					
Ala	Arg	Asn	Ser	Thr	Gly	Lys	Ala	Leu	Pro	Asp	Leu	Cys	Arg	Asn	Phe			
		435					440					445						
His	Cys	Lys	Arg	Gly	Lys	Val	Cys	Gln	Ala	Asp	Lys	Gln	Gly	Lys	Pro			
	450					455					460							
Ser	Cys	Ile	Cys	Gln	Asp	Pro	Ala	Ala	Cys	Pro	Ser	Thr	Lys	Asp	Tyr			
465					470					475					480			
Lys	Arg	Val	Cys	Gly	Thr	Asp	Asn	Lys	Thr	Tyr	Asp	Gly	Thr	Cys	Gln			
				485					490					495				
Leu	Phe	Gly	Thr	Lys	Cys	Gln	Leu	Glu	Gly	Thr	Lys	Met	Gly	Arg	Gln			
			500					505					510					
Leu	His	Leu	Asp	Tyr	Met	Gly	Ala	Cys	Lys	His	Ile	Pro	His	Cys	Thr			
		515				520						525						
Asp	Tyr	Glu	Val	Asn	Gln	Phe	Pro	Leu	Arg	Met	Arg	Asp	Trp	Leu	Lys			
	530					535					540							
Asn	Ile	Leu	Met	Gln	Tyr	Tyr	Glu	Arg	Asp	Gln	Asp	Thr	Ser	Ala	Phe			
545					550					555					560			
Leu	Thr	Glu	Lys	Gln	Arg	Asn	Lys	Val	Lys	Lys	Ile	Tyr	Leu	Asn	Glu			
				565					570					575				

Lys Arg Leu Val Ser Gly Glu His Pro Val Glu Leu Leu Leu His Asp
580 585 590
Phe Glu Lys Asn Tyr His Met Tyr Leu Tyr Pro Val His Trp Gln Phe
595 600 605
Tyr Gln Leu Asp Gln His Pro Val Asp Arg Ser Leu Thr His Ser Glu
610 615 620
Leu Ala Pro Leu Arg Ala Ser Leu Val Pro Met Glu His Cys Ile Thr
625 630 635 640
Arg Phe Phe Gln Glu Cys Asp Gly Asp Gln Asp Lys Leu Ile Thr Leu
645 650 655
Lys Glu Trp Cys His Cys Phe Ala Ile Lys Glu Glu Asp Ile Asn Glu
660 665 670
Asn Leu Leu Phe
675

<210> 37
<211> 434
<212> PRT
<213> *Homo sapiens*

5

<400> 37

Met Leu Pro Ala Arg Cys Ala Arg Leu Leu Thr Pro His Leu Leu Leu
1 5 10 15
Val Leu Val Gln Leu Ser Pro Ala Arg Gly His Arg Thr Thr Gly Pro
20 25 30
Arg Phe Leu Ile Ser Asp Arg Asp Pro Gln Cys Asn Leu His Cys Ser
35 40 45
Arg Thr Gln Pro Lys Pro Ile Cys Ala Ser Asp Gly Arg Ser Tyr Glu
50 55 60
Ser Met Cys Glu Tyr Gln Arg Ala Lys Cys Arg Asp Pro Thr Leu Gly
65 70 75 80
Val Val His Arg Gly Arg Cys Lys Asp Ala Gly Gln Ser Lys Cys Arg
85 90 95
Leu Glu Arg Ala Gln Ala Leu Glu Gln Ala Lys Lys Pro Gln Glu Ala
100 105 110
Val Phe Val Pro Glu Cys Gly Glu Asp Gly Ser Phe Thr Gln Val Gln
115 120 125
Cys His Thr Tyr Thr Gly Tyr Cys Trp Cys Val Thr Pro Asp Gly Lys
130 135 140
Pro Ile Ser Gly Ser Ser Val Gln Asn Lys Thr Pro Val Cys Ser Gly
145 150 155 160
Ser Val Thr Asp Lys Pro Leu Ser Gln Gly Asn Ser Gly Arg Lys Asp
165 170 175
Asp Gly Ser Lys Pro Thr Pro Thr Met Glu Thr Gln Pro Val Phe Asp

10

[illegible]

<210> 38
<211> 3669
<212> ADN
<213> *Homo sapiens*

<400> 38

gcctgctgcc	gcctgggccc	cgccgagcgg	agctagcgcc	gcgcgcagag	cacacgctcg	60
cgctccagct	cccctcctgc	gcggttcagt	actgtgtccc	ctgaccgcag	cctctgcgag	120
ccccgcgcgc	aggaccacgg	cccgtcccc	gccgccgcga	gggccccgag	cgaaggaagg	180

aagggaggcg cgctgtgcgc cccgcggagc ccgcgaaccc cgctcgctgc cggctgcccc	240
gcctggctgg caccatgctg cccgcgcgct gcgcccgcct gctcacgccc cacttgcctg	300
tggtgttggt gcagctgtcc cctgctcgcg gccaccgcac cacaggcccc aggtttctaa	360
taagtgaccg tgacccacag tgcaacctcc actgctccag gactcaaccc aaaccatct	420
gtgcctctga tggcaggctc tacgagtcca tgtgtgagta ccagcgagcc aagtgccgag	480
acccgaccct gggcgtggcg catcgaggta gatgcaaaga tgctggccag agcaagtgtc	540
gcctggagcg ggctcaagcc ctggagcaag ccaagaagcc tcaggaagct gtgtttgtcc	600
cagagtgtgg cgaggatggc tcctttaccc aggtgcagtg ccatacttac actgggtact	660
gctgggtgtg caccocggat gggaaagccc tcagtggctc ttctgtgcag aataaaactc	720
ctgtatgttc aggttcagtc accgacaagc ccttgagcca gggtaactca ggaaggaaag	780
atgacgggtc taagccgaca cccacgatgg agaccagcc ggtgttcgat ggagatgaaa	840
tcacagcccc aactctatgg attaaacact tggatgatcaa ggactccaaa ctgaacaaca	900
ccaacataag aaattcagag aaagtctatt cgtgtgacca ggagaggcag agtgccctgg	960
aagaggccca gcagaatccc cgtgagggtc ttgtcatccc tgaatgtgcc cctgggggac	1020
tctataagcc agtgcaatgc caccagtcca ctggctactg ctgggtgtgtg ctgggtggaca	1080
cagggcgccc gctgcctggg acctccacac gctacgtgat gccagttgt gagagcgacg	1140
ccagggccaa gactacagag gcggatgacc ccttcaagga caggagagta ccaggctgtc	1200
cagaaggga gaaaatggag tttatcacca gcctactgga tgctctcacc actgacatgg	1260
ttcaggccat taactcagca gcgcccactg gaggtgggag gttctcagag ccagacccca	1320
gccacaccct ggaggagcgg gtagtgcact ggtatttcag ccagctggac agcaatagca	1380
gcaacgacat taacaagcgg gagatgaagc ccttcaagcg ctacgtgaag aagaaagcca	1440
agcccaagaa atgtgcccg cgtttcaccc actactgtga cctgaacaaa gacaaggtca	1500
tttactgcc tgagctgaag ggctgcctgg gtgttagcaa agaaggacgc ctctctaaag	1560
gagcagaaaa cccaaggga ggtggagagt ccaggaggc aggatggatc accagacacc	1620
taaccttcag cgttgcccat ggcctgcga catcccgtgt aacataagt gtgcccacca	1680
tgtttgcaact ttaataact ctacttgcg tgttttgtt ttggtttcat tttaaaacac	1740
caatatctaa taccacagt ggaaaaggaa aggaagaaa gactttatc tctctcttat	1800
tgtaagtttt tggatctgct actgacaact tttagagggt tttggggggg tgggggagg	1860
tgttgttggg gcctgagaag aaagagattt atatgctgta tataaatata tatgtaaatt	1920
gtatagttct tttgtacagg cattggcatt gctgtttgtt tatttctctc cctctgcctg	1980
ctgtgggtgg tgggcactct ggacacatag tccagctttc taaaatccag gactctatcc	2040
tgggcctact aaacttctgt ttggagactg acccttgtgt ataaagacgg gagtccgtca	2100

attgtactgc ggactccacg agttcttttc tgggtgggagg actatattgc cccatgccat 2160
tagttgtcaa aattgataag tcaacttggtc ctccggccttg tccagggagg ttgggctaag 2220
gagagatgga aactgcctcg ggagaggaag ggagtccaga tcccatgaat agccacaca 2280
ggtaccggct ctccaggggt ccgtgcattc ctgctctccg gacccccaaa gggcccagca 2340
ttggtgggtg caccagtatc ttagtgacct tcggagcaaa ttatccacaa aggatttgca 2400
ttacgtcact cgaaacgttt tcatccatgc ttagcatcta ctctgtataa cgcagagag 2460
gggaggcaaa gaagaaaagg acacacggaa gggcctttaa aaaagtagat atttaatatc 2520
taagcagggg aggggacagg acagaaagcc tgcactgagg ggtgcggtgc caacagggaa 2580
actcttcacc tccctgcaaa cctaccagtg aggtccccag agacgcagct gtctcagtgc 2640
ccaggggcag attgggtgtg acctctccac tcctccatct cctgctgttg tcctagtggc 2700
tatcacaggc ctgggtgggt ggggtggggg aagtgtcagt caccttggtg gtaacactaa 2760
agttgttttg ttggtttttt aaaaacccaa tactgagggt ctctctgttc cctcaagttt 2820
tcttatgggc ttccaggctt taagctaatt ccagaagtaa aactgatctt gggtttccta 2880
ttctgcctcc cctagaaggg cagggtgat aaccagcta cagggaatc cgggccagc 2940
ttccacagg catcacaggc atcttccgct gattctaggg tgggctgcc agccttctgg 3000
tctgagggc agctccctct gccaggtgc tgtgcctatt caagtggcct tcaggcagag 3060
cagcaagtgg cccttagcgc cccttcccat aagcagctgt ggtggcagtg agggaggttg 3120
ggtagccctg gactggtccc ctctcagat cacccttgca aatctggcct catcttgtat 3180
tccaaccoga catccctaaa agtaacctca ccggtcccg gtctggaagg cgttggcacc 3240
acaagcactg tccctgtggg aggagcacia ccttctcggg acaggatctg atgggtctt 3300
gggctaagg aggtccctgc tgtcctggag aaagtcttag aggttatctc aggaatgact 3360
ggtggccctg cccaacgtg gaaaggtggg aaggaagcct tctccatta gccccaatga 3420
gagaactcaa cgtgccggag ctgagtgggc cttgcacgag aactggccc cactttcagg 3480
cctggaggaa gcatgcacac atggagacgg cgcctgcctg tagatgtttg gatcttcgag 3540
atctccccag gcatcttgtc tcccacagga tcgtgtgtgt aggtggtgtt gtgtggtttt 3600
cctttgtgaa ggagagaggg aaactatttg tagcttggtt tataaaaaat aaaaaatggg 3660
taaactcttg 3669

<210> 39
<211> 2133
<212> ADN
<213> *Homo sapiens*
<400> 39

cgaggagagcg cgetctgcct gccgcctgcc tgccctgccac tgagggttcc cagcaccatg	60
agggcctgga tcttctttct cctttgcctg gccgggaggg ccttggcagc ccctcagcaa	120
gaagccctgc ctgatgagac agaggtggtg gaagaaactg tggcagaggt gactgaggta	180
tctgtgggag ctaatcctgt ccaggtggaa gtaggagaat ttgatgatgg tgcagaggaa	240
accgaagagg aggtggtggc ggaaaatccc tgccagaacc accactgcaa acacggcaag	300
gtgtgcgagc tggatgagaa caacaccccc atgtgcgtgt gccaggaccc caccagctgc	360
ccagccccca ttggcgagtt tgagaagggtg tgcagcaatg acaacaagac cttcgactct	420
tcctgccact tctttgccac aaagtgcacc ctggagggca ccaagaaggc ccacaagctc	480
cacctggact acatcgggcc ttgcaaatac atccccctt gcctggactc tgagctgacc	540
gaattcccc tgccatgcg ggactggctc aagaacgtcc tggtcaccct gtatgagagg	600
gatgaggaca acaaccttct gactgagaag cagaagctgc gggatgaaga gatccatgag	660
aatgagaagc gcctggaggc aggagaccac ccctggagc tgctggcccg ggacttcgag	720
aagaactata acatgtacat cttccctgta cactggcagt tcggccagct ggaccagcac	780
cccattgacg ggtacctctc ccacaccgag ctggctccac tgcgtgctcc cctcatcccc	840
atggagcatt gcaccaccg ctttttcgag acctgtgacc tggacaatga caagtacatc	900
gcctggatg agtgggccg ctgcttcggc atcaagcaga aggatatcga caaggatctt	960
gtgatctaaa tccactcctt ccacagtacc ggattctctc tttaacctc cccttcgtgt	1020
ttcccccaat-gtttaaaatg-ttgggatggg-ttgttgttct-gcctggagac-aagggtgctaa	1080
catagattta agtgaatata ttaacggtgc taaaaatgaa aattctaacc caagacatga	1140
cattcttagc tgtaacttaa ctattaaggc cttttccaca cgcattaata gtcccatttt	1200
tctcttgcca tttgtagctt tgccattgt cttattggca catgggtgga cacggatctg	1260
ctgggctctg ccttaaacac acattgcagc ttcaactttt ctctttagtg ttctgtttga	1320
aactaatact taccgagtca gactttgtgt tcatttcatt tcagggtctt ggctgcctgt	1380
gggcttcccc aggtggcctg gaggtgggca aagggaagta acagacacac gatgttgtca	1440
aggatggttt tgggactaga ggctcagtgg tgggagagat ccctgcagaa tccaccaacc	1500
agaacgtggt ttgcctgagg ctgtaactga gagaaagatt ctggggctgt cttatgaaaa	1560
tatagacatt ctcacataag ccagttcat caccatttcc tcctttacct ttcagtgcag	1620
tttcttttca cattaggctg ttggttcaaa cttttgggag cacggactgt cagtctctg	1680
ggaagtggtc agcgatcct gcagggttc tcctcctctg tcttttgag aaccagggt	1740
cttctcaggg gctctaggga ctgccaggt gtttcagcca ggaaggccaa aatcaagagt	1800
gagatgtaga aagttgtaaa atagaaaaag tggagttggt gaatcgggtg ttctttcctc	1860

acatttggat gattgtcata aggttttttag catgttcctc cttttcttca ccctcccctt	1920
tgttcttcta ttaatcaaga gaaacttcaa agttaatggg atggtcggat ctcacaggct	1980
gagaactcgt tcacctcaa gcatttcattg aaaaagctgc ttcttattaa tcatacaaac	2040
tctcaccatg atgtgaagag tttcacaaat ctttcaaaat aaaaagtaat gacttagaaa	2100
ctgaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa	2133

<210> 40
 <211> 2678
 <212> ADN
 <213> *Coturnix coturnix*

<400> 40

gtcagacctt cctcctcaga agctcacaga aaaacacgct ttctgaaaga ttccacactc	60
aatgccaaaa tataccacag gaaaattttg caaggctcac ggatttccag tgcaccactg	120
gctaaccaag taggagcacc tcttctactg ccatgaaagg aaaccttcaa accctaccac	180
tgagccatta actaccatcc tgtttaagat ctgaaaaaca tgaagactgt attgctcctg	240
atttgtcttc taggatctgc tttcaccact ccaaccgatc cattgaaata ccaatttggg	300
gcccatggac agaaaactgc agagaagcat aaatatactc attctgaaat gccagaggaa	360
gagaacacag ggttttgtaa caaaggatgt gtgctgtctg gccacaggac cataaaagca	420
gaggtaccgg tactggatac acagaaggat gagccctggg cttccagaag acaaggacaa	480
ggtgatgggt agcatcaaac aaaaaacagc ctgaggagca ttaacttcct tactctgcac	540
agtaatccag ggttggcttc tgataaccag gaaagcaact ctggcagcag cagggaacag	600
cacagctctg agcaccacca gcccaggagg cacaggaaac acggcaacat ggctggccag	660
tgggctctga gaggagaaag tccagtggat gctcttggc tggttcgtga gcgcaacaca	720
tggaataaca ataaaaacac agttggccta gatgaaaaca acaatggaag tgaagaagag	780
gaagctgggg aggaagaaga tgaggaatgg ggtgaagaaa ctgattacag ggatatgaaa	840
cacagagccc gtgggacaag ccatggaaga gaatacagaa gatggcaaaa tgaaaacagc	900
cggccatctg gtgaattctt gagagattcc agtctgccag tacgtataac caagagacac	960
ggtgagaaat tcagcatgga ggaggaaagt caggaaaagc tctacaagga aggaaaactc	1020
cctctctcaa agaaaaatca taatgaggat caagggtgaa aaagacaaag tgaagaaagt	1080
aaagagcatt ttcaagtagt caatcagcgc aaacacagag cagtgcagaa aaggcaggat	1140
aaggagggca gcaatgctga ggaggatgat aatgatagtg gtgatgatgg tgaggaagat	1200
cttggcaatg tctggaggga agcagtctac gaggaagagg aaagaatgca aagcaatgac	1260
caggacagta tcaactaaca gcaaaaagag gaaataactg ctggagatga cagtggagtt	1320
tatagggaga tgcaggatta caaagggtgac aaaattaaag atgttactca ctctgaagac	1380

aatcattacc accatgagcc ccctaattcc agcagcaagc aacaactgca aacaagtagc	1440
tctgttgaga gcatgaattc aacagagcat gaggatgagg ttaagaccac aggaggttca	1500
tatcatgagg aaagtgcaag gaacagcact gggaaggctc tcccgatctt ttgtagaaac	1560
ttccactgca aaagaggaaa agtctgcca gcagacaagc aaggaaaacc cagctgtatt	1620
tgccaagatc ctgctgcttg cccttcacc aaagattata agcgtgtttg tggcactgat	1680
aataagactt acgatggtac gtgccaactc tttggcacca aatgtcaact tgaagggaca	1740
aaaatgggac gccagctgca cctggactat atgggtgcct gcaaacacat accccactgt	1800
actgattacg aagtgaatca gttccctctc cgtatgagag actggctcaa aaacatccta	1860
atgcaatatt atgaactga tcaggatacg tctgcatttc taaccgaaaa gcaaaggaat	1920
aaggtcaaaa agatatacct gaatgagaag cgtctcgtct ctggtgagca cccagttgag	1980
cttctcctgc atgactttga gaaaaactac cacatgtatc totatcctgt gcaactggca	2040
ttttatcagc ttgaccagca cccagttgac agatcactga ctcatcaga gctcgtcct	2100
ttgagagcct ccctcgttcc catggaacac tgcataacc gtttcttcca ggagtgtgat	2160
ggagaccaag acaacttat cactttgaaa gagtggtgcc actgctttgc gattaaggaa	2220
gaagacataa atgaaaatct cctgttctga gccacctga gcagaatccc catgcagcgc	2280
tacagcttgt caaacatgca atgcccattt atgactgcaa ttaacagctc tgtaatttt	2340
caggaataag ttggcataag attcttggag gcagaacaag tcgctcttgg ataacacaag	2400
tgctaattg ttacaaattc attaacagca gtagtgttta agagctctaa gtagctcata	2460
cttaagagt gtttccctct gcacgtacca ataactctct agtaagacga ctaacttgat	2520
gactgagttg ttcacaaaac cttccgtag aattatagga tgtggatttt ataatacacc	2580
gataaaaact actttgaaat aggttttctt ttcctgtcgt ttactgtcag tagctctctg	2640
catagaaatg tcaaataaac agatcttgtt ttggtttc	2678

<210> 41
 <211> 2808
 <212> ADN
 <213> *Homo sapiens*

<400> 41

cgcatgaga ggccagcctg ccagggaat ccaggaatct gcaacaaaa cgatgacagt	60
ctgaaatact ctctggtgcc aacctocaaa ttctcgtctg tcaactcaga cccccactag	120
ttgacagagc agcagaatat caactccagt agacttgaat gtgcctctgg gcaaagaagc	180
agagctaacg aggaaggga tttaaagagt ttttcttggg tgtttgtcaa acttttattc	240
cctgtctgtg tgcagagggg attcaacttc aattttctgc agtggctctg ggtccagccc	300
cttacttaaa gatctggaaa gcatgaagac tgggcctttt ttcctatgtc tcttgggaac	360

tgcagctgca atcccgacaa atgcaagatt atttatctgat cattccaaac caactgctga	420
aacggtagca cctgacaaca ctgcaatccc cagtttatgg gctgaagctg aagaaaatga	480
aaaagaaaca gcagtatcca cagaagacga ttcccacat aaggctgaaa aatcatcagt	540
actaaagtca aaagaggaaa gccatgaaca gtcagcagaa cagggcaaga gttctagcca	600
agagctggga ttgaaggatc aagaggacag tgatggtcac ttaagtgtga atttgagta	660
tgcaccaact gaaggtacat tggacataaa agaagatatg attgagcctc aggagaaaa	720
actctcagag aacactgatt ttttggtccc tgggtttagt tccttcacag attctaacca	780
acaagaaagt atcacaaga gagaggaaaa ccaagaacaa cctagaaatt attcacatca	840
tcagttgaac aggagcagta aacatagcca aggcctaagg gatcaaggaa accaagagca	900
ggatccaaat atttccaatg gagaagagga agaagaaaa gagccaggtg aagttggtac	960
ccacaatgat aaccaagaaa gaaagacaga attgcccagg gagcatgcta acagcaagca	1020
ggagggaagac aatacccaat ctgatgatat tttggaagag tctgatcaac caactcaagt	1080
aagcaagatg caggaggatg aatttgatca gggtaaccaa gaacaagaag ataactccaa	1140
tgcagaaatg gaagaggaaa atgcatcgaa cgtcaataag cacattcaag aaactgaatg	1200
gcagagtcaa gagggtaaaa ctggcctaga agctatcagc aaccacaaag agacagaaga	1260
aaagactgtt tctgaggctc tgctcatgga acctactgat gatggtaata ccacgcccag	1320
aaatcatgga gttgatgatg atggcgatga tgatggcgat gatggcgga ctgatggccc	1380
caggcacagt gcaagtgatg actacttcat cccaagccag gcctttctgg aggccgagag	1440
agctcaatcc attgcctatc acctcaaaat tgaggagcaa agagaaaaag tacatgaaaa	1500
tgaaaatata ggtaccactg agcctggaga gcaccaagag gccaaagaaag cagagaactc	1560
atcaaatgag gaggaaacgt caagtgaagg caacatgagg gtgcatgctg tggattcttg	1620
catgagcttc cagtgtaaaa gaggccacat ctgtaaggca gaccaacagg gaaaacctca	1680
ctgtgtctgc caggatccag tgacttgtcc tccaacaaaa ccccttgatc aagtttgtgg	1740
cactgacaat cagacctatg ctagtctctg tcacttattc gctactaaat gcagactgga	1800
ggggacaaa aaggggcata aactccagct ggattatttt ggagcctgca aatctattcc	1860
tacttgtacg gactttgaag tgattcagtt tcctctacgg atgagagact ggctcaagaa	1920
tatcctcatg cagctttatg aagccaactc tgaacatgct ggttatctaa atgagaagca	1980
gagaaataaa gtcaagaaaa ttacctgga tgaaaagagg cttttggctg gggaccatcc	2040
cattgatctt ctcttaaggg actttaagaa aaactaccac atgtatgtgt atcctgtgca	2100
ctggcagttt agtgaacttg accaacaccc tatggataga gtcttgacac attctgaact	2160
tgctctctctg cgagcatctc tggtgcccat ggaacactgc ataaccggtt tctttgagga	2220
gtgtgacccc aacaaggata agcacatcac cctgaaggag tggggccact gctttggaat	2280

taaagaagag gacatagatg aaaatctctt gttttgaacg aagattttta agaactcaac 2340
 tttccagcat cctcctctgt tctaaccact tcagaaatat atgcagctgt gatacttgta 2400
 gatttatatt tagcaaatg ttagcatgta tgacaagaca atgagagtaa ttgcttgaca 2460
 acaacctatg caccaggtat ttaacattaa ctttggaac aaaaatgtac aattaagtaa 2520
 agtcaacata tgcaaaatac tgtacattgt gaacagaagt ttaattcata gtaatttcac 2580
 tctctgcatt gacttatgag ataattaatg attaaactat taatgataaa aataatgcat 2640
 ttgtattgtt cataatatca tgtgcacttc aagaaaatgg aatgctactc ttttgtgggt 2700
 tacgtgtatt attttcaata tcttaatacc ctaataaaga gtccataaaa atccaaaaaa 2760
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2808

<210> 42

<211> 4804

<212> ADN

<213> *Homo sapiens*

<400> 42

agcggccaga cgctcggcgg cggcgtgtgg caggagcgca ggggcgcgag cgggcgatca 60
 gccttcccgg cgaccgtgcc ggggagctc gagcaactcg gactagggga cccgggccgg 120
 cccccaagat gccggcgatc gcgggtattgg cggcgccgc cggggcgtgg tgcttctcc 180
 aagtcgagag ccggcacctg gaocgcctcg ccggaggcgc ggcccccaac caccgcaatt 240
 tcctagacaa tgaccagtgg ctgagcaccg tctcccagta cgaccgggac aagtactgga 300
 accgctttcg agacgatgat tatttcagaa actggaatcc caacaagccc tttgaccaag 360
 ccctggaccc atccaaggac ccctgcctga aggtaaaatg cagccctcac aaagtgtgtg 420
 tgaccagga ctaccagacc gccctgtgtg tcagccgcaa gcacctgctc ccagggcaa 480
 agaaggggaa cgtggccag aaacactggg ttggaccttc gaatttggtc aagtgcgaagc 540
 cctgtcccggt ggacacgtca gccatggtct gcggctcaga tggccactcc tacacatcca 600
 agtgcgaatt ggagttccat gcttgttcta ctggcaaaag cctcgccacc ctctgtgatg 660
 ggccctgtcc ctgtctccca gagcctgagc caccaaagca caaggcagaa aggagtgcct 720
 gcacagacaa ggagttgcgg aaccttgcc cccggctgaa ggattggttt ggagctctcc 780
 acgaggatgc gaacagagtc atcaagccca ccagctccaa cacagcccaa ggcaggtttg 840
 acactagcat cctgcccatc tgcaaggact ccctgggctg gatgttcaac aagttggaca 900
 tgaactatga cctcctgctt gacccttcag agatcaatgc catctacctg gataagtaag 960
 agccctgtat caagcctctt ttcaactcgt gtgactcctt caaggatggc aagctttcta 1020
 acaatgagtg gtgctactgc ttccagaagc ctggaggtct cccttgccag aatgaaatga 1080
 acagaattca gaagctgagt aaggggaaaa gcctgttggg ggccctcata cctcgggtga 1140

atgaggaggg ctattacaaa gccacacagt gccacggcag cacggggcag tgctggtgtg	1200
tggacaaaata tgggaatgag ttggctggct ccaggaaaca ggggtgctgtg agctgtgaag	1260
aggagcagga aacctcaggg gatthttggca gtggtgggtc cgtggtcctg ctggatgacc	1320
tagaatatga acgggagctg ggaccaaagg acaagagggg gaagctgagg gtgcacaccc	1380
gagccgtgac agaggatgat gaggatgagg atgatgacaa agaggatgag gtcgggtaca	1440
tatggtagtg cccacaagaa agaggacaca agthtttgac aaaattgcaa gtcacttcct	1500
attcctgcat ttgtatctaa gactccaagg caccaaggtc tcttctccat tgttgcctc	1560
tatacccgac ctaaggthtg gaagacaact gcttgttccc agaggattct gatthttgcat	1620
atgthttgat gggagaaagg gtgtgtgtt thttthtttg ttgtgtthta thttthtgat	1680
agggaagtca ttggcttaat tagagcctcc thcctthctg tgagaththt ccaacaagca	1740
tgtgathttac gtggaattct gacagtgcag ggagcccca cctctthaaa tgtcaaagac	1800
cctthttgat taccacact ggtggttatt acagcatggt tcccagcctt acagtgtcta	1860
agtgcctctc ttgtgtcctg tagatgttgt gaaaagaaa aaacaaaaa atacaccaca	1920
ctgtacttht thccctgcc cccgtactg ccggtgatta thattaaaa thagthththt	1980
tcacatcatt atatctggct tcctataaac aacagcctta attcagtcaa gactccctth	2040
gggaattcat thtattaaaa attggtgtct ggatacttcc ctgtacatgc ataaatatgc	2100
atgcatgtac aggaagactg tatgtgtgtg ccttgcacac acaccatac ctctcagaaa	2160
aagtgggtat attaaaaact cgaaaaacaa tgataaattt ctgagcttgt ccagacctgg	2220
aaacaaattt ctggaataag aaattthtat taaagtctt thttgacta acagthggct	2280
cttgtagcct gcaggctgag gaagtctctt ctctgtgcat cagcagagtt actgaaagcc	2340
tctgattgag aaaaaacctc cgtctgcta aatcacttht ctgcagaag ccatgcgact	2400
cccacacgac acgggcagct tcacaagcca tctctthcat thctgcttga agccccctg	2460
gctgcagcaa tcctgtctgc cataggttht thccttctt acctactcaa gggtththt	2520
taaggcatgc acacatatct cctgttctct gagagtacca tgggtthctt taaaagaaga	2580
aaattthcta thctgaactc aatgththtg thttactccc thtctactga caaatcatga	2640
taagggcaca aaagctgtac agathththt thttaaccac tcaatccaa atggaggcct	2700
acaagaaca tcgtaataac acatggaagc aaacctggg ththtaagag caaattctgt	2760
ccccctca cthccccaa tgacaagata ctaatgaaga aagthcttca ccatagtgtt	2820
tgtthtaact aaactcattg gagtctagtt ccaaattthg tagggtcatc atctctacat	2880
thcttaggat thctctccct atcaagctgg ccagataca agtaccaaac agtagtctct	2940
gaagthcccc catttcttc agtaccagtc tataagctac tgtccgccac tgaththcat	3000

ctatcagggt gtcctaata caaatcagcca cccaagcaag cctctctggc ccacatatct	3060
atctcttgcc ttccccatg aacttcagcc tgtccacaca aaagccacat aaactcaagc	3120
aagaaatatg ttcagccaaa acatgattat agtggcagct gaccaatccc ccatcccaaa	3180
ttgaccattt agatgtacca actcacctta aattagcatg ttccaatcca gtcggcattg	3240
cctgaataca gtagcatcat acctatagtt ggtottagat aagaaatgaa ctacttgata	3300
tagcaaaagtc ctttggttc gtaaataacc ctgaggtttt gtacttactt tccccatag	3360
aagacagacc ataggcaaac tctgttttg gatctcaact ccatcacctt tgtttcaata	3420
ttttttttct ctcttgaaca aaactgagat aatttagaaa acagggtgctt aattgcaata	3480
aaattactat gaagtatat aaaaatcacg acattgtaaa atctcacttt agatcatcaa	3540
agaaaacat tgttactatc tcctttgagc ttaggaaaat gtacaagaga acaaattaaa	3600
attgaaaaat tgatttcact tagaaaaact tctaggaaca ggggtgaacca ctgattttta	3660
tttgcctaata tatcttatga caagtatcaa attaagatga cacttaaaga tccttagcat	3720
taacttaatg atggagaaga gtgctcaaca gacagttccc agtaaggtaa tgagatgcca	3780
ttttccgaga cattctaaga agatattttg attcattaaa acattaaata aaaagccctc	3840
ctcagattgg aacccccaaa tcgatggagt cacattaata atacttttca tgccctcactt	3900
tgacatgaca gcattcgatt tttttaaga tctttaatac tttccatgag tactaaagat	3960
tgtaatgagt tacottatcc ttagaagtag aatgtttgct ttcttcttct tggaaatagt	4020
cctccaaaaa gtccacttgt accagtgaca agaagtcact tgtatagtga ccaagtacac	4080
ataaaaatacc gattataaaa atattgtaat aaaccacttc ctcatctgat actggtatta	4140
agctgaatcc tcactaattc acttttgaaa agttcttaat gaacagtttt attcttatac	4200
catgaacagc tcttttaata caatttcctt ggtatccaac ttaaattcca gaattttgtt	4260
ccaccatggt tagttatttg ccatatgaat gtattcttgt ttctaccaat attctaggca	4320
tgagaatacc ttaatactga gttggagatt ttgtatctgt ttctcctcct ttacgccccat	4380
tttctaccc acagcagcaa atgacaacgt gtctgtccag gtctgtcccc ctgctcatcc	4440
caggatgcca ctacatctt tttcttctt gttacccttg acccagctg tacaagtaac	4500
atccaagagc ccattctaca agtgggtggt tttggtctt ttataacttt ttctcaaagt	4560
cactgatggt tgttcctggt aaatgtatag cattgtaatg agagcccatc aaatcctgag	4620
tgtcagtttg ttgtccctat tgtagatgaa atagtgatgt agcaaaaacc tagtaaatc	4680
tgaatgcttt tccacgtaga cttatctgga atgtgaacac aactctttgg ttaatagtaa	4740
atgcttaact gtagtcctga gtaggtgcat ttctgtctgt ctcaataaat tttactttgt	4800
ctgc	4804

<210> 43
 <211> 1370
 <212> ADN
 <213> *Rattus norvegicus*

<400> 43

```

ctggcctcca actcactgct tccatcctgc ccagtgtcct ctcgagtccc ggacccgagc      60
acgatgtgga aacgctggct ggcgctcgcg ctggtgacca tcgccctggt ccacggcgag      120
gaggaacaaa gaagcaaata caagatctgc gccaatgtgt tttgtggagc tggccgggaa      180
tgcccgctca cggagaaggg ggagccaacg tgcctctgca ttgagcaatg caaacctcac      240
aagaggcctg tgtgtggcag taatggcaag acctacctca accattgtga acttcacaga      300
gacgcctgcc tcaactggatc caagatccag gttgattatg atgggcactg caaagaaaag      360
aagtctgtga gtccatccgc cagccccggt gtctgtatc aggctaaccg tgatgagctg      420
cggcgccgga tcatccagtg gctggaagcc gagatcattc cagatggctg gttctctaaa      480
ggcagtaact acagtgagat cctagacaag tactttaaga gctttgataa tggtgactct      540
cacctggact ccagcgaatt cctgaaatc gtggagcaga atgaaacagc cgtcaacatc      600
accgcttacc ccaatcagga gaacaacaaa ctgctcagag gcctctgtgt tgatgccctc      660
attgaactgt ccgatgagaa cgctgactgg aaactcagct tccaagagtt cctcaagtgc      720
ctcaaccat ccttcaacc tcttgagaag aagtgcgcc tggaggacga aacctatgca      780
gatggagctg agaccgaggt ggactgcaat cgctgtgtct gttcctgtgg aactgggtc      840
tgcacagcga tgacctgtga tggaaagaat cagaaggggg tccagacca cacagaggag      900
gagatgacga gatatgcca ggaactccag aagcaccagg gaacagcaga aaagaccaag      960
aaggtgaaca ccaaagagat ctaagaagag gcacgtagca cctcatctgg aaccagcac     1020
ctcctcttca gcgctaagcc cagtatacag cgtctgtggc aatcaccgaa tcaccagtat     1080
ttgcttgtac ggcagcaaat cttatctgtt tgttttgcaa taaaggaagt gaggggtggt     1140
ggctagccag ggcaggcagg ccacaacttt cacttctagg aatgctttaa gagacactaa     1200
agggcacctt ggggcaggag gcgagtatcc ggttggcaga ggagcagagg caggtctgaa     1260
tgaaaccttt ctggggtcag ctgtgaggat acaacaggaa aagcatgtga tgttaggggg     1320
aacactgagc tggccctgct ggaggaaata gggggagctt ggtggggagg     1370

```

<210> 44
 <211> 303
 <212> PRT
 <213> *Homo sapiens*

<400> 44

```

Met Arg Ala Trp Ile Phe Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu
1           5           10          15

```

Ala Ala Pro Gln Gln Glu Ala Leu Pro Asp Glu Thr Glu Val Val Glu
20 25 30
Glu Thr Val Ala Glu Val Thr Glu Val Ser Val Gly Ala Asn Pro Val
35 40 45
Gln Val Glu Val Gly Glu Phe Asp Asp Gly Ala Glu Glu Thr Glu Glu
50 55 60
Glu Val Val Ala Glu Asn Pro Cys Gln Asn His His Cys Lys His Gly
65 70 75 80
Lys Val Cys Glu Leu Asp Glu Asn Asn Thr Pro Met Cys Val Cys Gln
85 90 95
Asp Pro Thr Ser Cys Pro Ala Pro Ile Gly Glu Phe Glu Lys Val Cys
100 105 110
Ser Asn Asp Asn Lys Thr Phe Asp Ser Ser Cys His Phe Phe Ala Thr
115 120 125
Lys Cys Thr Leu Glu Gly Thr Lys Lys Gly His Lys Leu His Leu Asp
130 135 140
Tyr Ile Gly Pro Cys Lys Tyr Ile Pro Pro Cys Leu Asp Ser Glu Leu
145 150 155 160
Thr Glu Phe Pro Leu Arg Met Arg Asp Trp Leu Lys Asn Val Leu Val
165 170 175
Thr Leu Tyr Glu Arg Asp Glu Asp Asn Asn Leu Leu Thr Glu Lys Gln
180 185 190
Lys Leu Arg Val Lys Lys Ile His Glu Asn Glu Lys Arg Leu Glu Ala
195 200 205
Gly Asp His Pro Val Glu Leu Leu Ala Arg Asp Phe Glu Lys Asn Tyr
210 215 220
Asn Met Tyr Ile Phe Pro Val His Trp Gln Phe Gly Gln Leu Asp Gln
225 230 235 240
His Pro Ile Asp Gly Tyr Leu Ser His Thr Glu Leu Ala Pro Leu Arg
245 250 255
Ala Pro Leu Ile Pro Met Glu His Cys Thr Thr Arg Phe Phe Glu Thr
260 265 270
Cys Asp Leu Asp Asn Asp Lys Tyr Ile Ala Leu Asp Glu Trp Ala Gly
275 280 285
Cys Phe Gly Ile Lys Gln Lys Asp Ile Asp Lys Asp Leu Val Ile
290 295 300

<210> 45
<211> 439
5 <212> PRT
<213> *Homo sapiens*
<400> 45

Met Pro Ala Ile Ala Val Leu Ala Ala Ala Ala Ala Ala Trp Cys Phe

1	5	10	15
Leu Gln Val Glu Ser Arg His Leu Asp Ala Leu Ala Gly Gly Ala Gly	20	25	30
Pro Asn His Gly Asn Phe Leu Asp Asn Asp Gln Trp Leu Ser Thr Val	35	40	45
Ser Gln Tyr Asp Arg Asp Lys Tyr Trp Asn Arg Phe Arg Asp Asp Asp	50	55	60
Tyr Phe Arg Asn Trp Asn Pro Asn Lys Pro Phe Asp Gln Ala Leu Asp	65	70	75
Pro Ser Lys Asp Pro Cys Leu Lys Val Lys Cys Ser Pro His Lys Val	85	90	95
Cys Val Thr Gln Asp Tyr Gln Thr Ala Leu Cys Val Ser Arg Lys His	100	105	110
Leu Leu Pro Arg Gln Lys Lys Gly Asn Val Ala Gln Lys His Trp Val	115	120	125
Gly Pro Ser Asn Leu Val Lys Cys Lys Pro Cys Pro Val Ala Gln Ser	130	135	140
Ala Met Val Cys Gly Ser Asp Gly His Ser Tyr Thr Ser Lys Cys Lys	145	150	155
Leu Glu Phe His Ala Cys Ser Thr Gly Lys Ser Leu Ala Thr Leu Cys	165	170	175
Asp Gly Pro Cys Pro Cys Leu Pro Glu Pro Glu Pro Pro Lys His Lys	180	185	190
Ala Glu Arg Ser Ala Cys Thr Asp Lys Glu Leu Arg Asn Leu Ala Ser	195	200	205
Arg Leu Lys Asp Trp Phe Gly Ala Leu His Glu Asp Ala Asn Arg Val	210	215	220
Ile Lys Pro Thr Ser Ser Asn Thr Ala Gln Gly Arg Phe Asp Thr Ser	225	230	235
Ile Leu Pro Ile Cys Lys Asp Ser Leu Gly Trp Met Phe Asn Lys Leu	245	250	255
Asp Met Asn Tyr Asp Leu Leu Leu Asp Pro Ser Glu Ile Asn Ala Ile	260	265	270
Tyr Leu Asp Lys Tyr Glu Pro Cys Ile Lys Pro Leu Phe Asn Ser Cys	275	280	285
Asp Ser Phe Lys Asp Gly Lys Leu Ser Asn Asn Glu Trp Cys Tyr Cys	290	295	300
Phe Gln Lys Pro Gly Gly Leu Pro Cys Gln Asn Glu Met Asn Arg Ile	305	310	315
Gln Lys Leu Ser Lys Gly Lys Ser Leu Leu Gly Ala Phe Ile Pro Arg	325	330	335
Cys Asn Glu Glu Gly Tyr Tyr Lys Ala Thr Gln Cys His Gly Ser Thr			

340										345					350						
Gly	Gln	Cys	Trp	Cys	Val	Asp	Lys	Tyr	Gly	Asn	Glu	Leu	Ala	Gly	Ser						
355								360							365						
Arg	Lys	Gln	Gly	Ala	Val	Ser	Cys	Glu	Glu	Glu	Gln	Glu	Thr	Ser	Gly						
370							375							380							
Asp	Phe	Gly	Ser	Gly	Gly	Ser	Val	Val	Leu	Leu	Asp	Asp	Leu	Glu	Tyr						
385							390							395							
Glu	Arg	Glu	Leu	Gly	Pro	Lys	Asp	Lys	Glu	Gly	Lys	Leu	Arg	Val	His						
			405							410							415				
Thr	Arg	Ala	Val	Thr	Glu	Asp	Asp	Glu	Asp	Glu	Asp	Asp	Asp	Lys	Glu						
			420							425							430				
Asp	Glu	Val	Gly	Tyr	Ile	Trp															
435																					

<210> 46
<211> 664
<212> PRT
<213> *Homo sapiens*
<400> 46

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

Gln Asp Pro Asn Ile Ser Asn Gly Glu Glu Glu Glu Glu Lys Glu Pro
195 200 205

Gly Glu Val Gly Thr His Asn Asp Asn Gln Glu Arg Lys Thr Glu Leu
210 215 220

Pro Arg Glu His Ala Asn Ser Lys Gln Glu Glu Asp Asn Thr Gln Ser
225 230 235 240

Asp Asp Ile Leu Glu Glu Ser Asp Gln Pro Thr Gln Val Ser Lys Met
245 250 255

Gln Glu Asp Glu Phe Asp Gln Gly Asn Gln Glu Gln Glu Asp Asn Ser
260 265 270

Asn Ala Glu Met Glu Glu Glu Asn Ala Ser Asn Val Asn Lys His Ile
275 280 285

Gln Glu Thr Glu Trp Gln Ser Gln Glu Gly Lys Thr Gly Leu Glu Ala
290 295 300

Ile Ser Asn His Lys Glu Thr Glu Glu Lys Thr Val Ser Glu Ala Leu
305 310 315 320

Leu Met Glu Pro Thr Asp Asp Gly Asn Thr Thr Pro Arg Asn His Gly
325 330 335

Val Asp Asp Asp Gly Asp Asp Asp Gly Asp Asp Gly Gly Thr Asp Gly
340 345 350

Pro Arg His Ser Ala Ser Asp Asp Tyr Phe Ile Pro Ser Gln Ala Phe
355 360 365

Leu Glu Ala Glu Arg Ala Gln Ser Ile Ala Tyr His Leu Lys Ile Glu
370 375 380

Glu Gln Arg Glu Lys Val His Glu Asn Glu Asn Ile Gly Thr Thr Glu
385 390 395 400

Pro Gly Glu His Gln Glu Ala Lys Lys Ala Glu Asn Ser Ser Asn Glu
405 410 415

Glu Glu Thr Ser Ser Glu Gly Asn Met Arg Val His Ala Val Asp Ser
420 425 430

Cys Met Ser Phe Gln Cys Lys Arg Gly His Ile Cys Lys Ala Asp Gln
435 440 445

Gln Gly Lys Pro His Cys Val Cys Gln Asp Pro Val Thr Cys Pro Pro
450 455 460

Thr Lys Pro Leu Asp Gln Val Cys Gly Thr Asp Asn Gln Thr Tyr Ala
465 470 475 480

Ser Ser Cys His Leu Phe Ala Thr Lys Cys Arg Leu Glu Gly Thr Lys
485 490 495

Lys Gly His Gln Leu Gln Leu Asp Tyr Phe Gly Ala Cys Lys Ser Ile
500 505 510

Pro Thr Cys Thr Asp Phe Glu Val Ile Gln Phe Pro Leu Arg Met Arg
515 520 525

Asp Trp Leu Lys Asn Ile Leu Met Gln Leu Tyr Glu Ala Asn Ser Glu
530 535 540

His Ala Gly Tyr Leu Asn Glu Lys Gln Arg Asn Lys Val Lys Lys Ile
545 550 555 560

Tyr Leu Asp Glu Lys Arg Leu Leu Ala Gly Asp His Pro Ile Asp Leu
565 570 575

Leu Leu Arg Asp Phe Lys Lys Asn Tyr His Met Tyr Val Tyr Pro Val
580 585 590

His Trp Gln Phe Ser Glu Leu Asp Gln His Pro Met Asp Arg Val Leu
595 600 605

Thr His Ser Glu Leu Ala Pro Leu Arg Ala Ser Leu Val Pro Met Glu
610 615 620

His Cys Ile Thr Arg Phe Phe Glu Glu Cys Asp Pro Asn Lys Asp Lys
625 630 635 640

His Ile Thr Leu Lys Glu Trp Gly His Cys Phe Gly Ile Lys Glu Glu
645 650 655

Asp Ile Asp Glu Asn Leu Leu Phe
660

<210> 47
<211> 306
<212> PRT
<213> *Rattus norvegicus*

5

<400> 47

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

His Gly Glu Glu Glu Gln Arg Ser Lys Ser Lys Ile Cys Ala Asn Val
20 25 30

Phe Cys Gly Ala Gly Arg Glu Cys Ala Val Thr Glu Lys Gly Glu Pro
35 40 45

Thr Cys Leu Cys Ile Glu Gln Cys Lys Pro His Lys Arg Pro Val Cys
50 55 60

Gly Ser Asn Gly Lys Thr Tyr Leu Asn His Cys Glu Leu His Arg Asp
65 70 75 80

Ala Cys Leu Thr Gly Ser Lys Ile Gln Val Asp Tyr Asp Gly His Cys
85 90 95

Lys Glu Lys Lys Ser Val Ser Pro Ser Ala Ser Pro Val Val Cys Tyr
100 105 110

Gln Ala Asn Arg Asp Glu Leu Arg Arg Arg Ile Ile Gln Trp Leu Glu
115 120 125

Ala Glu Ile Ile Pro Asp Gly Trp Phe Ser Lys Gly Ser Asn Tyr Ser
130 135 140

10

Glu Ile Leu Asp Lys Tyr Phe Lys Ser Phe Asp Asn Gly Asp Ser His
 145 150 155 160
 Leu Asp Ser Ser Glu Phe Leu Lys Phe Val Glu Gln Asn Glu Thr Ala
 165 170 175
 Val Asn Ile Thr Ala Tyr Pro Asn Gln Glu Asn Asn Lys Leu Leu Arg
 180 185 190
 Gly Leu Cys Val Asp Ala Leu Ile Glu Leu Ser Asp Glu Asn Ala Asp
 195 200 205
 Trp Lys Leu Ser Phe Gln Glu Phe Leu Lys Cys Leu Asn Pro Ser Phe
 210 215 220
 Asn Pro Pro Glu Lys Lys Cys Ala Leu Glu Asp Glu Thr Tyr Ala Asp
 225 230 235 240
 Gly Ala Glu Thr Glu Val Asp Cys Asn Arg Cys Val Cys Ser Cys Gly
 245 250 255
 His Trp Val Cys Thr Ala Met Thr Cys Asp Gly Lys Asn Gln Lys Gly
 260 265 270
 Val Gln Thr His Thr Glu Glu Glu Met Thr Arg Tyr Ala Gln Glu Leu
 275 280 285
 Gln Lys His Gln Gly Thr Ala Glu Lys Thr Lys Lys Val Asn Thr Lys
 290 295 300
 Glu Ile
 305

<210> 48
 <211> 676
 <212> PRT
 <213> *Coturnix coturnix*

<400> 48

Met Lys Thr Val Leu Leu Leu Ile Cys Leu Leu Gly Ser Ala Phe Thr
 1 5 10 15
 Thr Pro Thr Asp Pro Leu Asn Tyr Gln Phe Gly Ala His Gly Gln Lys
 20 25 30
 Thr Ala Glu Lys His Lys Tyr Thr His Ser Glu Met Pro Glu Glu Glu
 35 40 45
 Asn Thr Gly Phe Val Asn Lys Gly Asp Val Leu Ser Gly His Arg Thr
 50 55 60
 Ile Lys Ala Glu Val Pro Val Leu Asp Thr Gln Lys Asp Glu Pro Trp
 65 70 75 80
 Ala Ser Arg Arg Gln Gly Gln Gly Asp Gly Glu His Gln Thr Lys Asn
 85 90 95
 Ser Leu Arg Ser Ile Asn Phe Leu Thr Leu His Ser Asn Pro Gly Leu
 100 105 110
 Ala Ser Asp Asn Gln Glu Ser Asn Ser Gly Ser Ser Arg Glu Gln His

115					120					125					
Ser	Ser	Glu	His	His	Gln	Pro	Arg	Arg	His	Arg	Lys	His	Gly	Asn	Met
130						135					140				
Ala	Gly	Gln	Trp	Ala	Leu	Arg	Gly	Glu	Ser	Pro	Val	Asp	Ala	Leu	Gly
145					150					155					160
Leu	Val	Arg	Glu	Arg	Asn	Thr	Trp	Lys	Tyr	Asn	Lys	Asn	Thr	Val	Gly
				165					170					175	
Leu	Asp	Glu	Asn	Asn	Asn	Gly	Ser	Glu	Glu	Glu	Glu	Ala	Gly	Glu	Glu
			180					185					190		
Glu	Asp	Glu	Glu	Trp	Gly	Glu	Glu	Thr	Asp	Tyr	Arg	Asp	Met	Lys	His
		195					200					205			
Arg	Ala	Arg	Gly	Thr	Ser	His	Gly	Arg	Glu	Tyr	Arg	Arg	Trp	Gln	Asn
	210					215					220				
Glu	Asn	Ser	Arg	Pro	Ser	Gly	Glu	Phe	Leu	Arg	Asp	Ser	Ser	Leu	Pro
225					230					235					240
Val	Arg	Ile	Thr	Lys	Arg	His	Gly	Glu	Lys	Phe	Ser	Met	Glu	Glu	Glu
				245					250					255	
Ser	Gln	Glu	Lys	Leu	Tyr	Lys	Glu	Gly	Lys	Leu	Pro	Leu	Ser	Lys	Lys
			260					265					270		
Asn	His	Asn	Glu	Asp	Gln	Gly	Glu	Lys	Arg	Gln	Ser	Glu	Glu	Ser	Lys
			275				280					285			
Glu	His	Phe	Gln	Val	Val	Asn	Gln	Arg	Lys	His	Arg	Ala	Val	Thr	Lys
	290					295					300				
Arg	Gln	Asp	Lys	Glu	Gly	Ser	Asn	Ala	Glu	Glu	Asp	Asp	Asn	Asp	Ser
305					310					315					320
Gly	Asp	Asp	Gly	Glu	Glu	Asp	Leu	Gly	Asn	Val	Trp	Arg	Glu	Ala	Val
			325						330					335	
Tyr	Glu	Glu	Glu	Glu	Arg	Met	Gln	Ser	Asn	Asp	Gln	Asp	Ser	Ile	Thr
			340					345					350		
Asn	Lys	Gln	Lys	Glu	Glu	Ile	Thr	Ala	Gly	Asp	Asp	Ser	Gly	Val	Tyr
		355					360					365			
Arg	Glu	Met	Gln	Asp	Tyr	Lys	Gly	Asp	Lys	Ile	Lys	Asp	Val	Thr	His
	370					375					380				
Ser	Glu	Asp	Asn	His	Tyr	His	His	Glu	Pro	Pro	Asn	Ser	Ser	Ser	Lys
385					390					395					400
Gln	Gln	Leu	Gln	Thr	Ser	Ser	Ser	Val	Glu	Ser	Met	Asn	Ser	Thr	Glu
				405						410				415	
His	Glu	Asp	Glu	Val	Lys	Thr	Thr	Gly	Gly	Ser	Tyr	His	Glu	Glu	Ser
			420					425					430		
Ala	Arg	Asn	Ser	Thr	Gly	Lys	Ala	Leu	Pro	Asp	Leu	Cys	Arg	Asn	Phe
		435					440					445			
His	Cys	Lys	Arg	Gly	Lys	Val	Cys	Gln	Ala	Asp	Lys	Gln	Gly	Lys	Pro

450	455	460
Ser Cys Ile Cys Gln Asp	Pro Ala Ala Cys Pro	Ser Thr Lys Asp Tyr
465	470	475 480
Lys Arg Val Cys Gly Thr	Asp Asn Lys Thr Tyr	Asp Gly Thr Cys Gln
	485	490 495
Leu Phe Gly Thr Lys Cys	Gln Leu Glu Gly Thr	Lys Met Gly Arg Gln
	500	505 510
Leu His Leu Asp Tyr Met	Gly Ala Cys Lys His	Ile Pro His Cys Thr
	515	520 525
Asp Tyr Glu Val Asn Gln	Phe Pro Leu Arg Met	Arg Asp Trp Leu Lys
	530	535 540
Asn Ile Leu Met Gln Tyr	Tyr Glu Arg Asp Gln	Asp Thr Ser Ala Phe
	545	550 555 560
Leu Thr Glu Lys Gln Arg	Asn Lys Val Lys Lys	Ile Tyr Leu Asn Glu
	565	570 575
Lys Arg Leu Val Ser Gly	Glu His Pro Val Glu	Leu Leu Leu His Asp
	580	585 590
Phe Glu Lys Asn Tyr His	Met Tyr Leu Tyr Pro	Val His Trp Gln Phe
	595	600 605
Tyr Gln Leu Asp Gln His	Pro Val Asp Arg Ser	Leu Thr His Ser Glu
	610	615 620
Leu Ala Pro Leu Arg Ala	Ser Leu Val Pro Met	Glu His Cys Ile Thr
	625	630 635 640
Arg Phe Phe Gln Glu Cys	Asp Gly Asp Gln Asp	Lys Leu Ile Thr Leu
	645	650 655
Lys Glu Trp Cys His Cys	Phe Ala Ile Lys Glu	Glu Asp Ile Asn Glu
	660	665 670
Asn Leu Leu Phe		
	675	

<210> 49
 <211> 4
 <212> PRT
 <213> Desconocido

<220>
 <223> Dominio conservado

<220>
 <221> DOMINIO
 <222> (1) .. (4)
 <223> Dominio conservado

<220>
 <221> misc_feature
 <222> (2) .. (2)
 <223> Xaa puede ser cualquier aminoácido de origen natural

<400> 49

Arg Xaa Lys Arg
 1

<210> 50

<211> 12
 <212> ADN
 <213> Desconocido,
 5 <220>
 <223> Secuencia de nucleótidos para el dominio conservado
 <220>
 <221> misc_feature
 10 <222> (1)..(12)
 <223> Secuencia de nucleótidos para el dominio conservado
 <220>
 <221> misc_feature
 15 <222> (4)..(6)
 <223> n es a, c, g, o t
 <400> 50
 20 cgcnnnaagc gc 12
 <210> 51
 <211> 6
 <212> PRT
 <213> Desconocido,
 25 <220>
 <223> Dominio conservado
 <220>
 30 <221> DOMINIO
 <222> (1) .. (6)
 <223> Dominio conservado
 <400> 51
 35 **Pro Leu Gly Leu Trp Ala**
 1 5
 <210> 52
 <211> 18
 40 <212> ADN
 <213> Desconocido,
 <220>
 <223> Secuencia de nucleótidos para el dominio conservado
 45 <220>
 <221> misc_feature
 <222> (i)..(18)
 <223> Secuencia de nucleótidos para el dominio conservado
 50 <400> 52
 ccctgggcc tgtgggcc 18
 <210> 53
 <211> 6
 55 <212> PRT
 <213> Desconocido,
 <220>
 60 <223> Dominio conservado
 <220>
 <221> DOMINIO
 <222> (1)..(6)

<223> Dominio conservado

<400> 53

5 Tyr Glu Val Asp Gly Trp
1 5

<210> 54
<211> 18
<212> ADN
10 <213> Desconocido,

<220>
<223> Secuencia de nucleótidos para el dominio conservado

15 <220>
<221> misc_feature
<222> (1)..(18)
<223> Secuencia de nucleótidos para el dominio conservado

20 <400> 54
tacgaggtgg acggctgg 18

<210> 55
<211> 7
25 <212> PRT
<213> Desconocido,

<220>
<223> Dominio conservado

30 <220>
<221> DOMINIO
<222> (1) .. (7)
<223> Dominio conservado

35 <400> 55

Val Asp Val Ala Asp Gly Trp
1 5

40 <210> 56
<211> 21
<212> ADN
<213> Desconocido,

45 <220>
<223> Secuencia de nucleótidos para el dominio conservado

<220>
<221> misc_feature
50 <222> (1)..(21)
<223> Secuencia de nucleótidos para el dominio conservado

<400> 56
gtggacgtgg ccgacggctg g 21

55 <210> 57
<211> 7
<212> PRT
<213> Desconocido,

60 <220>
<223> Dominio conservado

<220>

<221> DOMINIO
<222> (1) .. (7)
<223> Dominio conservado

5 <400> 57

Val Asp Gln Met Asp Gly Trp
1 5

10	$\langle 210 \rangle$ 58 $\langle 211 \rangle$ 21 $\langle 212 \rangle$ ADN $\langle 213 \rangle$ Desconocido,
----	---

15 <223> Secuencia de nucleótidos para el dominio conservado

20 <220>
<221> misc_feature
<222> (1)..(21)
<223> Secuencia de nucleótidos para el dominio conservado

<400> 58
gtggaccaga tggacggctg g 21

25 <210> 59
 <211> 6
 <212> PRT
 <213> Desconocido,

30 <220>
<223> Secuencia conservada

35 <220>
 <221> DOMINIO
 <222> (1)..(6)
 <223> Dominio conservado

<400> 59

Leu Glu Val Asp Gly Trp
1 5

45 <210> 60
 <211> 18
 <212> ADN
 <213> Desconocido.

<220>
<223> Secuencia de nucleótidos para el dominio conservado

50 <220>
 <221> misc_feature
 <222> (1) .. (18)
 <223> Secuencia de nucleótidos para el dominio conservado

55 <400> 60
ctggaggtgg acggctgg 18

	<210> 61
	<211> 6
60	<212> PRT
	<213> Desconocido.

<220>

<223> Dominio conservado
 <220>
 <221> DOMINIO
 5 <222> (1) .. (6)
 <223> Dominio conservado
 <400> 61
 Val Gln Val Asp Gly Trp
 10 1 5
 <210> 62
 <211> 18
 <212> ADN
 15 <213> Desconocido,
 <220>
 <223> Secuencia de nucleótidos para el dominio conservado
 20 <220>
 <221> misc_feature
 <222> (1)..(18)
 <223> Secuencia de nucleótidos para el dominio conservado
 25 <400> 62
 gtgcaggctgg acggctgg 18
 <210> 63
 <211> 7
 30 <212> PRT
 <213> Desconocido,
 <220>
 <223> Dominio conservado
 35 <220>
 <221> DOMINIO
 <222> (1) .. (7)
 <223> Dominio conservado
 40 <400> 63
 Val Asp Gln Val Asp Gly Trp
 1 5
 45 <210> 64
 <211> 21
 <212> ADN
 <213> Desconocido,
 50 <220>
 <223> Secuencia de nucleótidos para el dominio conservado
 <220>
 <221> misc_feature
 55 <222> (1) . (21)
 <223> Secuencia de nucleótidos para el dominio conservado
 <400> 64
 gtggaccagg tggacggctg g 21
 60 <210> 65
 <211> 4
 <212> PRT
 <213> Desconocido,

<220>
 <223> Dominio conservado

 5 <220>
 <221> DOMINIO
 <222> (1)..(4)
 <223> Dominio conservado

 10 <220>
 <221> misc_feature
 <222> (2)..(3)
 <223> Xaa puede ser cualquier aminoácido de origen natural

 15 <400> 65
 Asp Xaa Xaa Asp
 1

 20 <210> 66
 <211> 12
 <212> ADN
 <213> Desconocido,

 25 <220>
 <223> Secuencia de nucleótidos para el dominio conservado

 30 <220>
 <221> misc_feature
 <222> (1)..(12)
 <223> Secuencia de nucleótidos para el dominio conservado

 35 <220>
 <221> misc_feature
 <222> (4)..(9)
 <223> n es a, c, g, o t

 <400> 66
 gacnnnnnng ac 12

 40 <210> 67
 <211> 4
 <212> PRT
 <213> Desconocido,

 45 <220>
 <223> Dominio conservado

 50 <220>
 <221> DOMINIO
 <222> (1)..(4)
 <223> Dominio conservado

 <400> 67

 Arg Gly Leu Thr
 1
 55 <210> 68
 <211> 12
 <212> ADN
 <213> Desconocido,

 60 <220>
 <223> Secuencia de nucleótidos para el dominio conservado
 <220>
 <221> misc_feature

<222> (1)..(12)
 <223> Secuencia de nucleótidos para el dominio conservado

5 <400> 68
 cgcggcctga cc 12

<210> 69
 <211> 17
 <212> PRT
 10 <213> Desconocido,

<220>
 <223> Dominio conservado

15 <220>
 <221> DOMINIO
 <222> (1)..(17)
 <223> Dominio conservado

20 <400> 69

	Asn	Phe	Asp	Leu	Leu	Lys	Leu	Ala	Gly	Asp	Val	Glu	Ser	Asn	Pro	Gly
	1				5					10					15	
	Pro															

<210> 70
 25 <211> 51
 <212> ADN
 <213> Desconocido,

<220>
 30 <223> Secuencia de nucleótidos para el dominio conservado

<220>
 <221> misc_feature
 <222> (1)..(51)
 35 <223> Secuencia de nucleótidos para el dominio conservado

<400> 70
 aacttcgacc tgctgaagct ggcggcgac gtggagagca acccggccc c 51

40 <210> 71
 <211> 4
 <212> PRT
 <213> Desconocido,

45 <220>
 <223> Dominio conservado

<220>
 <221> DOMINIO
 50 <222> (1)..(4)
 <223> Dominio conservado

<220>
 <221> misc_feature
 55 <222> (2)..(2)
 <223> Xaa puede ser cualquier aminoácido de origen natural

<220>
 <221> misc_feature
 60 <222> (4)..(4)
 <223> Xaa puede ser cualquier aminoácido de origen natural
 <400> 71

Ala Xaa Ala Xaa**1**

5 <210> 72
 <211> 12
 <212> ADN
 <213> Desconocido,

 <220>
 <223> Secuencia de nucleótidos para el dominio conservado
 10
 <220>
 <221> misc_feature
 <222> (1)..(12)
 <223> Secuencia de nucleótidos para el dominio conservado
 15
 <220>
 <221> misc_feature
 <222> (4)..(6)
 <223> n es a, c, g, o t
 20
 <220>
 <221> misc_feature
 <222> (10)..(12)
 <223> n es a, c, g, o t
 25
 <400> 72
 gccnnngccn nn 12

 30 <210> 73
 <211> 3
 <212> PRT
 <213> Desconocido,

 <220>
 35 <223> Dominio conservado

 <220>
 <221> DOMINIO
 <222> (1)..(3)
 40 <223> Dominio conservado

 <400> 73

Leu Thr Lys
1
 45
 <210> 74
 <211> 9
 <212> ADN
 <213> Desconocido,
 50
 <220>
 <223> Secuencia de nucleótidos para el dominio conservado

 <220>
 55 <221> misc_feature
 <222> (1)..(9)
 <223> Secuencia de nucleótidos para el dominio conservado

 <400> 74
 60 ctgaccaag 9

 <210> 75
 <211> 9

<212> PRT
 <213> Desconocido,
 <220>
 5 <223> Secuencia conservada
 <220>
 <221> DOMINIO
 <222> (1)..(9)
 10 <223> Dominio conservado
 <400> 75
 Gly Gly Phe Leu Arg Lys Val Gly Gln
 1 5
 15 <210> 76
 <211> 27
 <212> ADN
 <213> Desconocido,
 20 <220>
 <223> Secuencia de nucleótidos para el dominio conservado
 <220>
 25 <221> misc_feature
 <222> (1)..(27)
 <223> Secuencia de nucleótidos para el dominio conservado
 <400> 76
 30 ggcggcttc tgcgcaaggt gggccag 27
 <210> 77
 <211> 17
 <212> PRT
 35 <213> Desconocido,
 <220>
 <223> Dominio conservado
 40 <220>
 <221> DOMINIO
 <222> (1)..(17)
 <223> Dominio conservado
 45 <400> 77
 Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly
 1 5 10 15
 Pro
 <210> 78
 <211> 23
 <212> ADN
 50 <213> Secuencia artificial
 <220>
 <223> Cebador sintético
 55 <220>
 <221> misc_feature
 <222> (1)..(23)
 <223> Cebador sintético
 60 <400> 78
 cgaagaggag gtggtggcgg aaa 23

5	<210> 79	
	<211> 26	
	<212> ADN	
	<213> Secuencia artificial	
10	<220>	
	<221> misc_feature	
	<222> (1)..(26)	
	<223> Cebador sintético	
15	<400> 79	
	ggtgtgtgtc ctcacccctc tcatac	26
	<210> 80	
	<211> 24	
20	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador sintético	
25	<220>	
	<221> misc_feature	
	<222> (1)..(24)	
	<223> Cebador sintético	
30	<400> 80	
	ctctctgtct ctcctgttcg acag	24
	<210> 81	
	<211> 24	
35	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador sintético	
40	<220>	
	<221> misc_feature	
	<222> (1)..(24)	
	<223> Cebador sintético	
45	<400> 81	
	aggggtctta ctcctggag gccca	24

REIVINDICACIONES

1. Uso de un polinucleótido de la familia de SPARC que codifica un polipéptido de la familia de SPARC con la secuencia de SEC ID N°: 1 en la preparación de un medicamento para aumentar la sensibilidad o para reducir la resistencia de un mamífero al que se le ha diagnosticado cáncer a un tratamiento terapéutico para el cáncer, en el que dicho mamífero muestra resistencia a dicho tratamiento terapéutico.
2. El uso de acuerdo con la reivindicación 1, en el que dicho tratamiento terapéutico es quimioterapia.
3. El uso de acuerdo con una cualquiera de las reivindicaciones 1-2, en el que el polinucleótido de la familia de SPARC comprende la SEC ID N°: 18.
4. El uso de acuerdo con una cualquiera de las reivindicaciones 1-3, en el que el cáncer es cáncer colorrectal.
5. El uso de acuerdo con una cualquiera de las reivindicaciones 2-4, en el que el agente quimioterapéutico es fluorouracilo.
6. Un polinucleótido de la familia de SPARC que codifica un polipéptido de la familia de SPARC con la secuencia de SEC ID N°: 1 para su uso para aumentar la sensibilidad o para reducir la resistencia de un mamífero al que se le ha diagnosticado cáncer a un tratamiento terapéutico para el cáncer, en el que dicho mamífero muestra resistencia a dicho tratamiento terapéutico.
7. Un polinucleótido de la familia de SPARC para su uso de acuerdo con la reivindicación 6, en el que dicho tratamiento terapéutico es quimioterapia.
8. Un polinucleótido de la familia de SPARC para su uso de acuerdo con una cualquiera de las reivindicaciones 6-7, en el que el polinucleótido de la familia de SPARC comprende la SEC ID N°: 18.
9. Un polinucleótido de la familia de SPARC para su uso de acuerdo con una cualquiera de las reivindicaciones 6-8, en el que el cáncer es cáncer colorrectal.
10. Un polinucleótido de la familia de SPARC para su uso de acuerdo con una cualquiera de las reivindicaciones 6-9, en el que el agente quimioterapéutico es fluorouracilo.

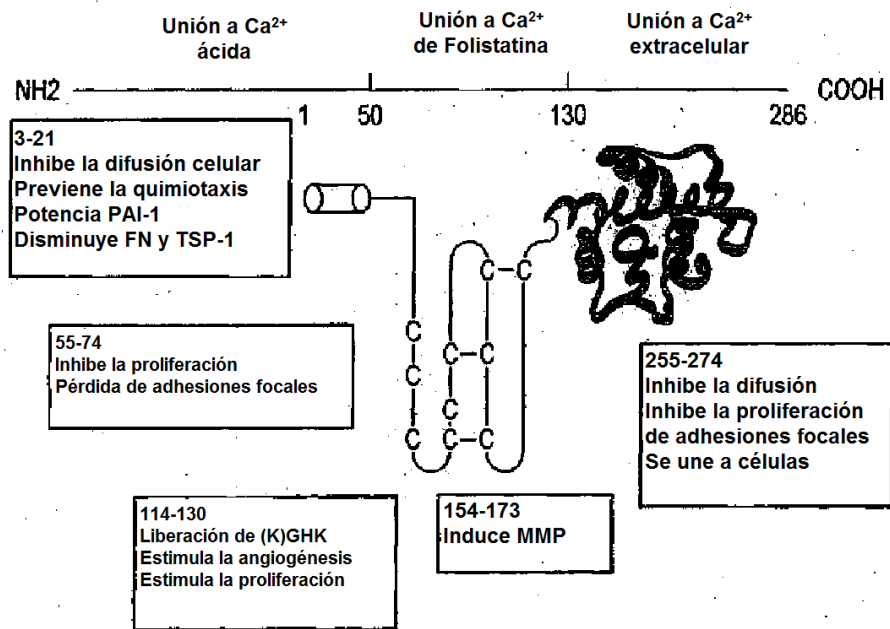


FIG. 1

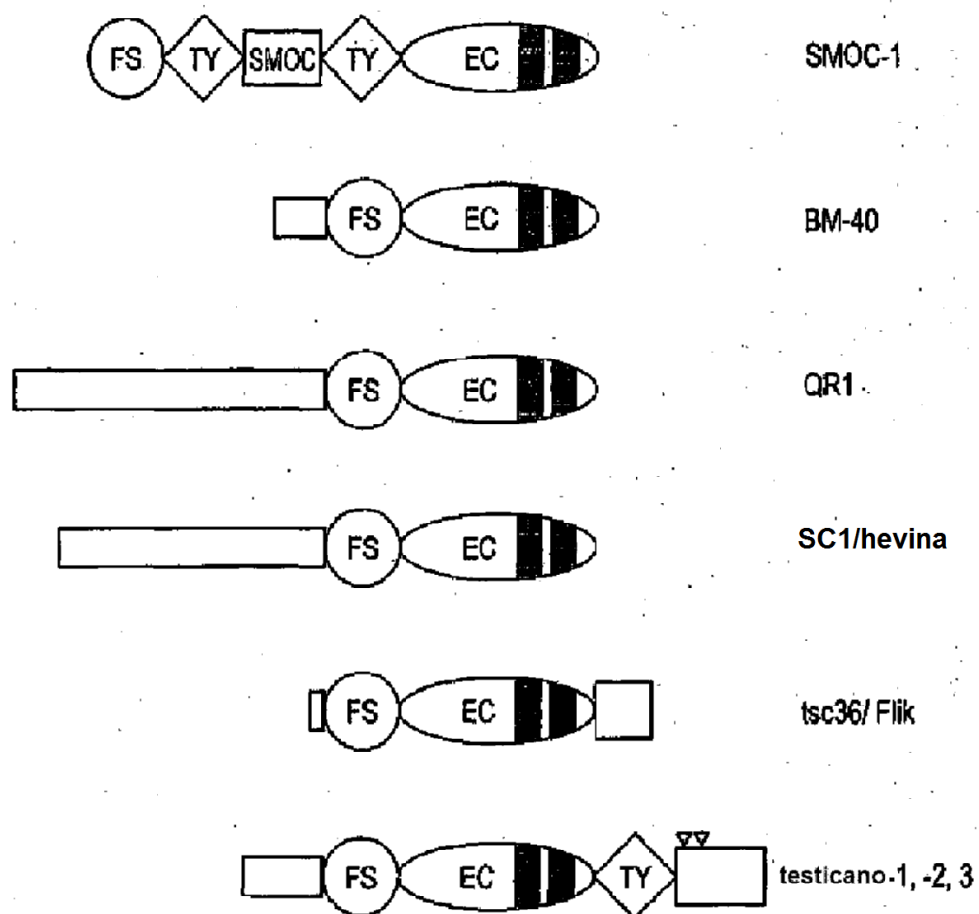


FIG. 2

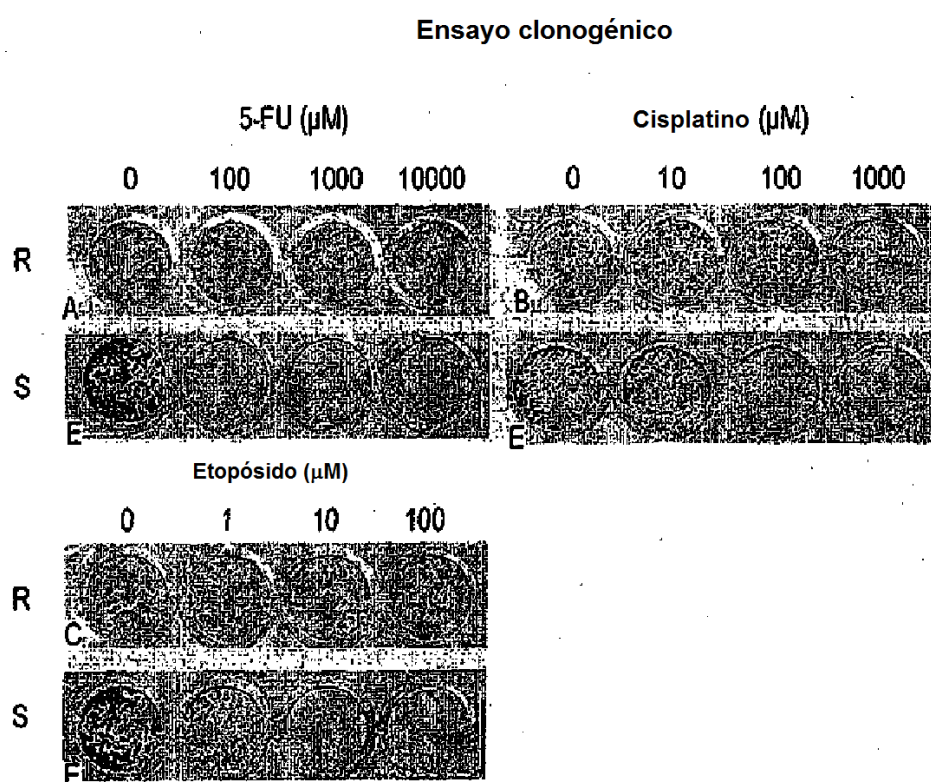


FIG. 3

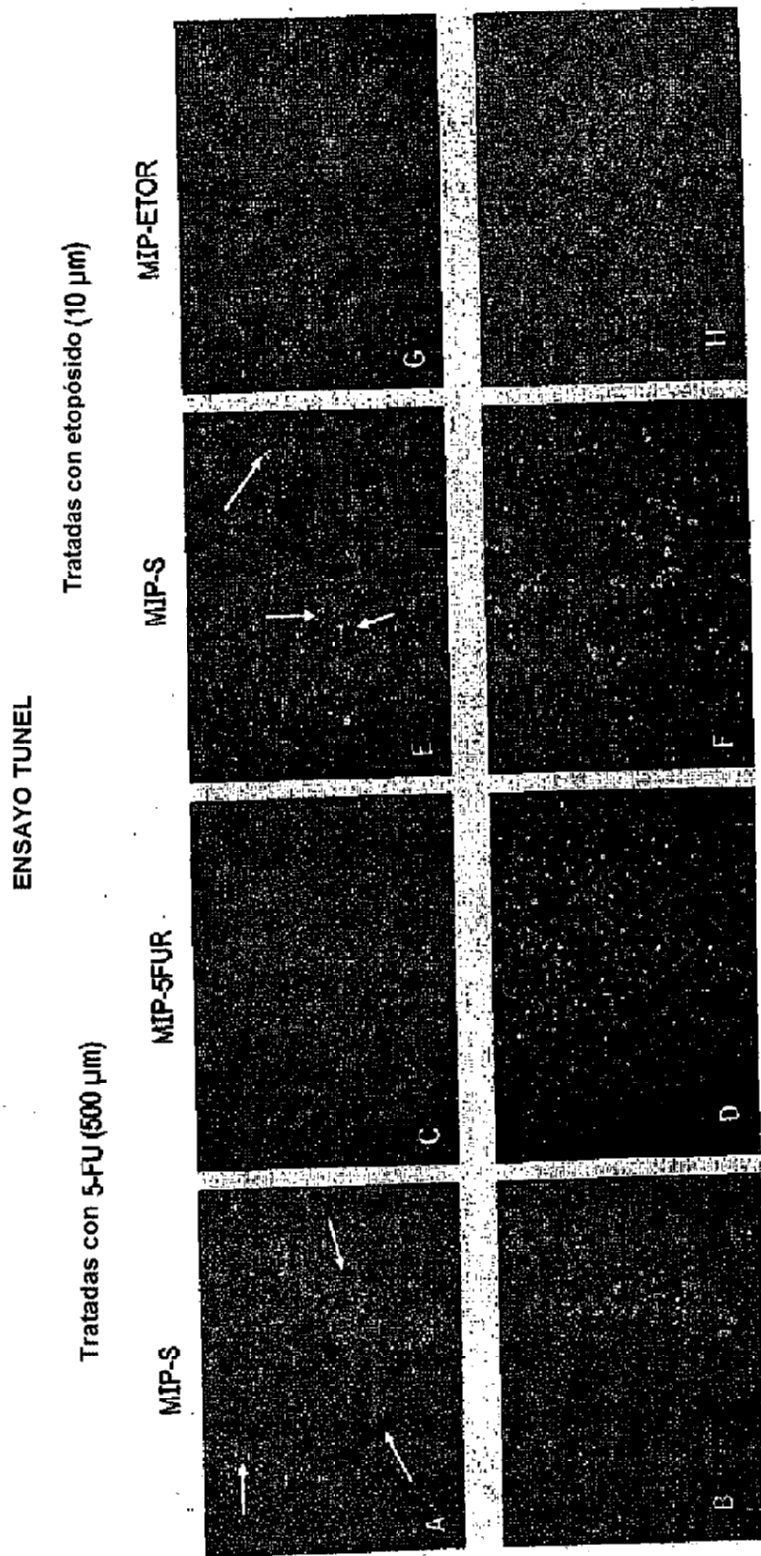


FIG. 4

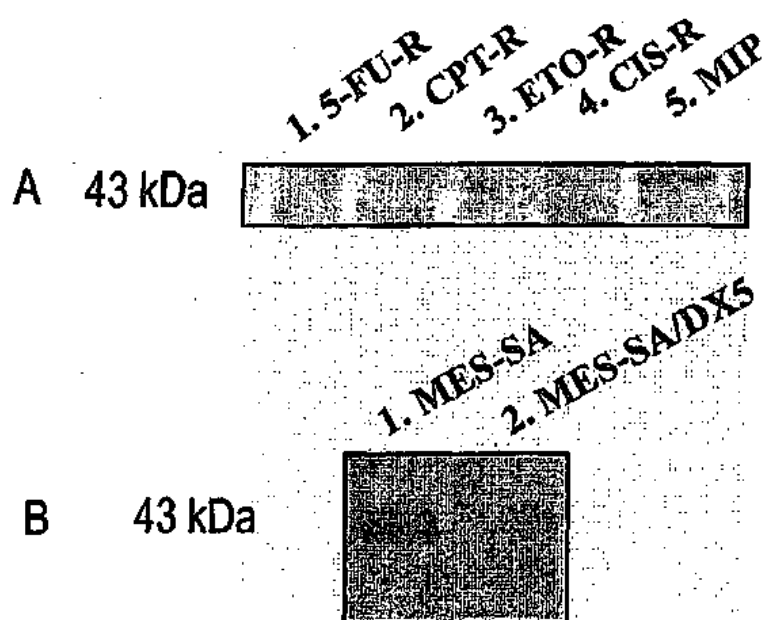


FIG. 5

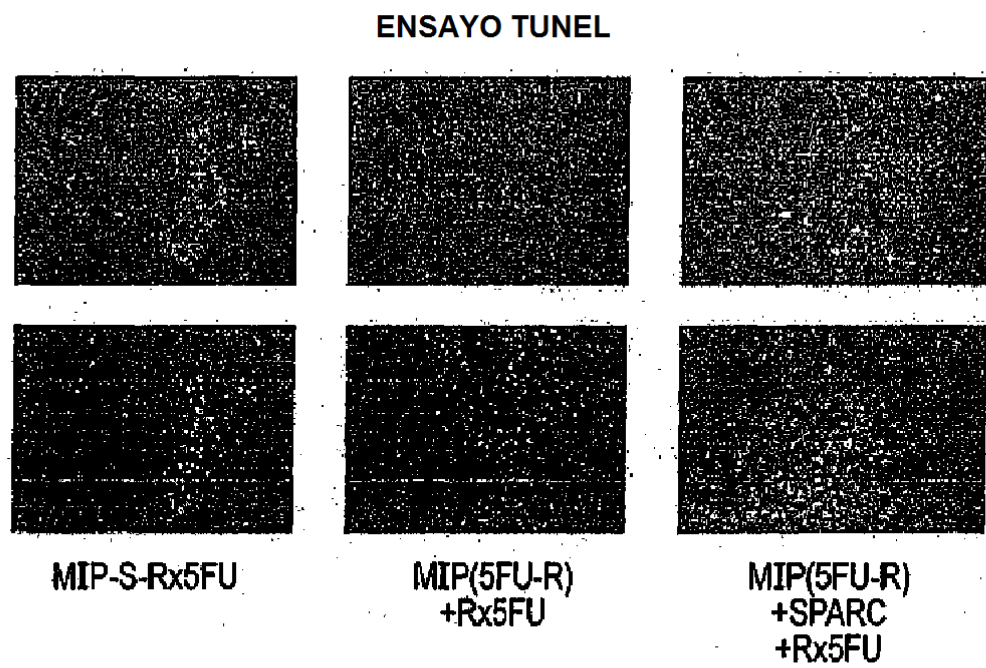


FIG. 6

EXPRESIÓN DE PROTEÍNA SPARC

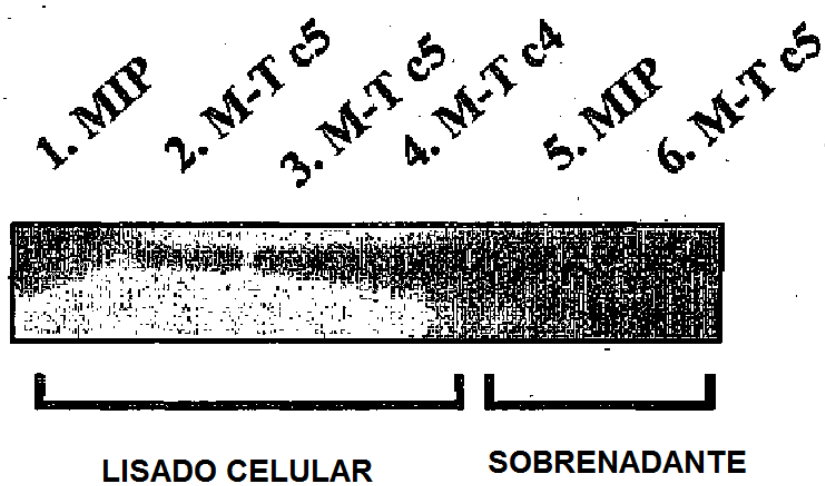


FIG. 7

EFFECTO DE LA SOBREEXPRESIÓN DE SPARC EN LA APOPTOSIS

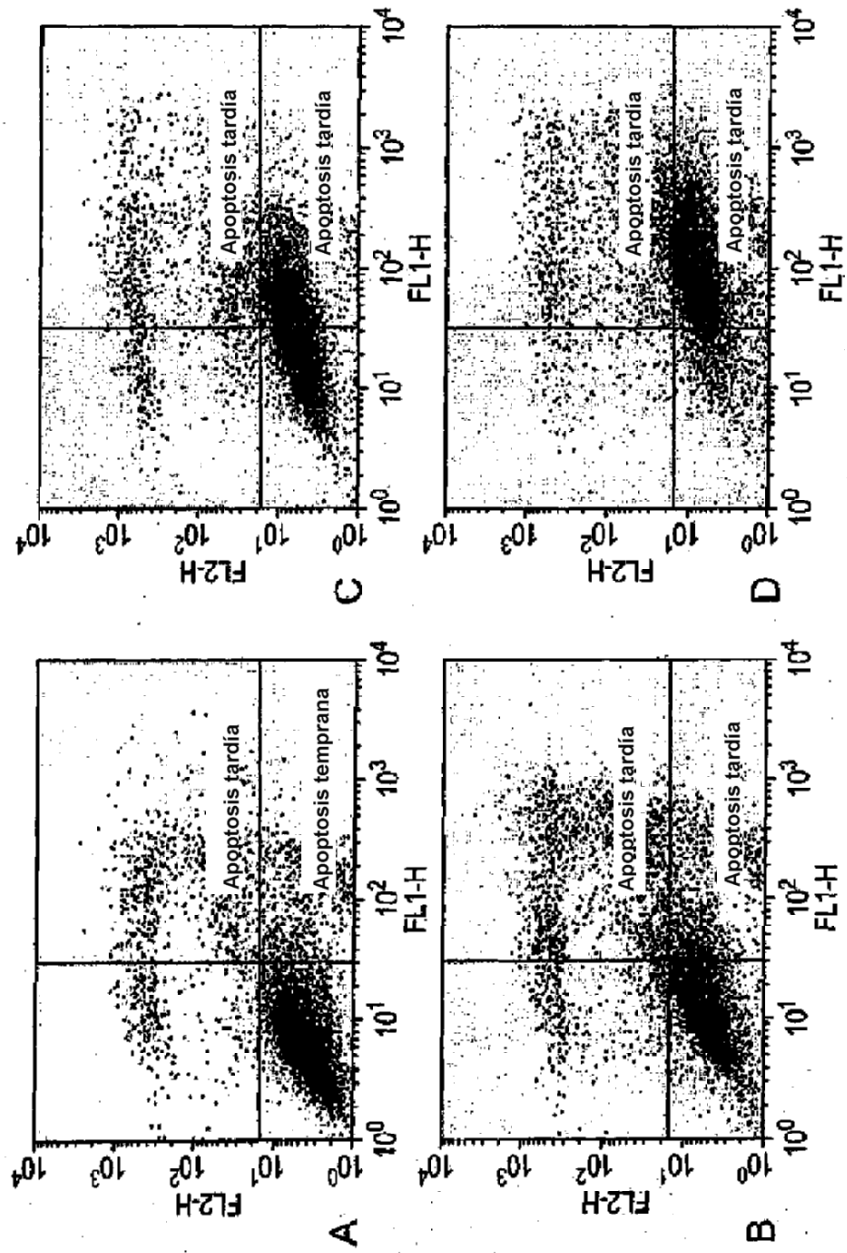


FIG. 8

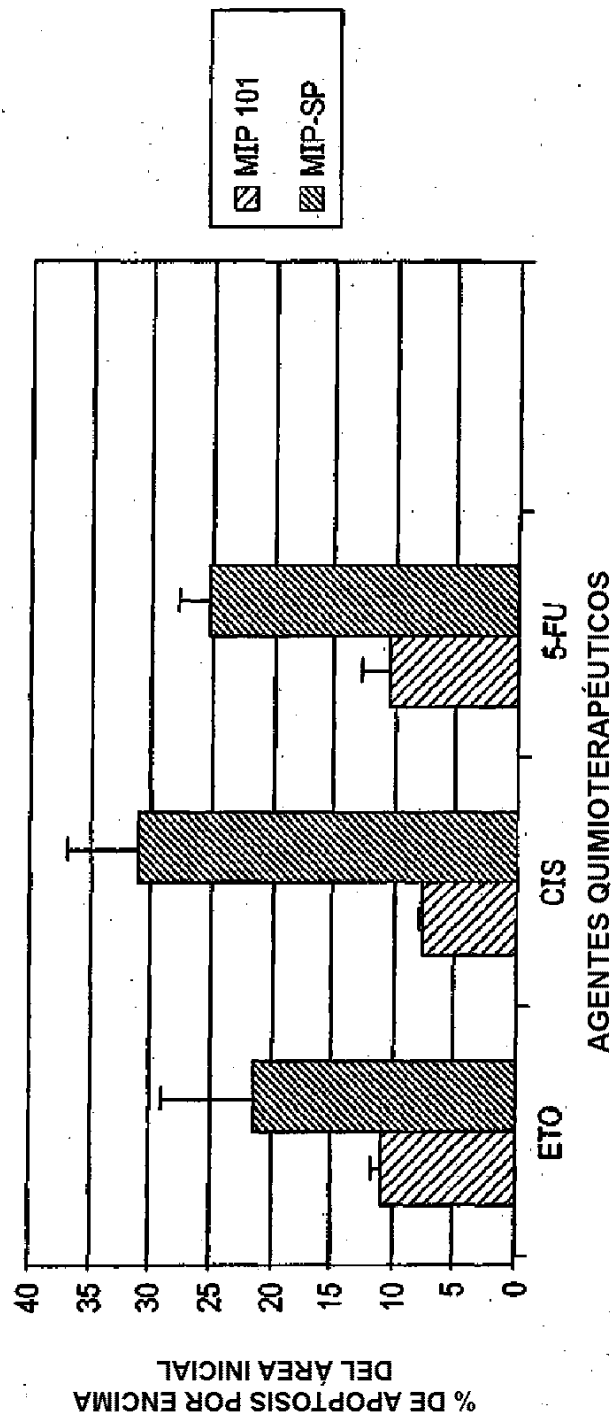


FIG. 9

RESPUESTA DE LOS TRANSFECTANTES DE SPARC A LA
QUIMIOTERAPIA

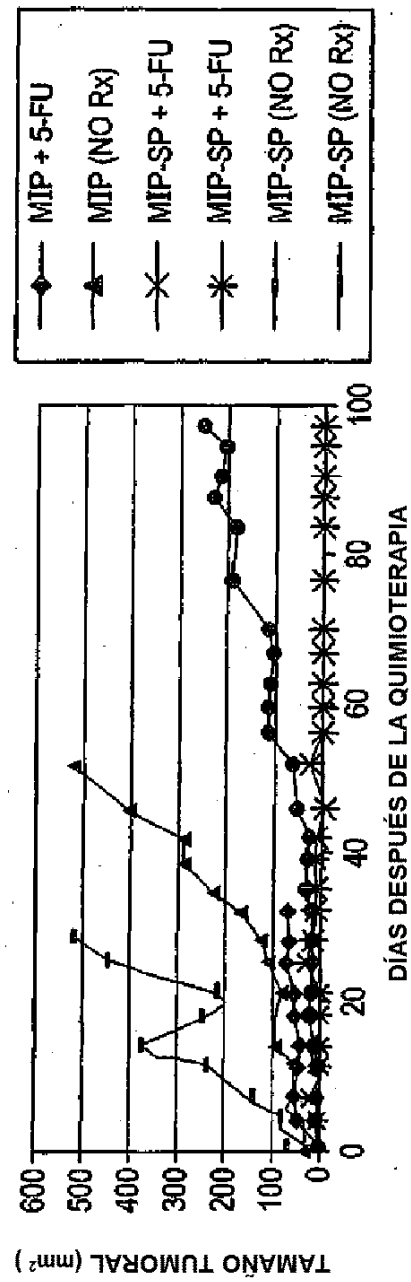


FIG. 10

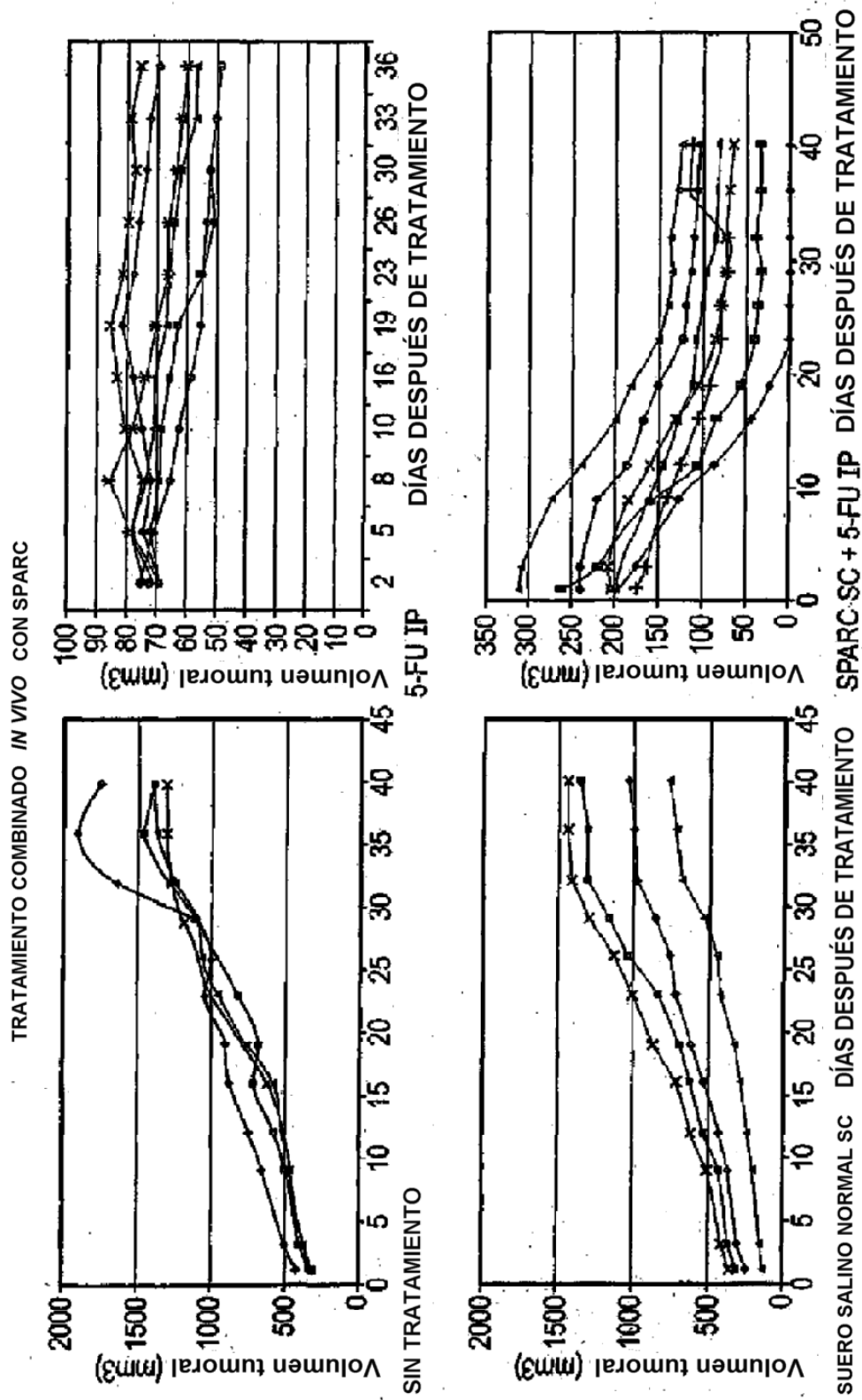
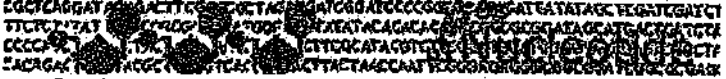


FIG. 11

[PubMed](#) [Nucleótido](#) [Proteína](#) [Genoma](#) [Estructura](#) [PMC](#) [Taxonomía](#) [OMIM](#) [Boo](#)

Buscar para

[Limites](#) [Vista previa/índice](#) [Historia](#) [Portapapeles](#) [Detalles](#)

1: NM_003118. Homo sapiens secr... [gi:4507170] [Links](#)

LOCUS	SPARC-	2133 pb	ARNm	lineal	PRI 05-NOV-2002
DEFINICIÓN	Proteína secretada de Homo sapiens, ácida, rica en cisteína (osteonectina) (SPARC), ARNm.				
REFERENCIA	NM_003118				
VERSIÓN	NM_003118.1 GI:4507170				
PAL. CLAVE	.				
FUENTE	Homo sapiens (humano)				
ORGANISMO	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCIA	1 (bases 1 a 2133)				
AUTORES	Termine, J.D., Kleinman, H.K., Whitson, S.W., Conn, X.M., McGarvey, M.L. y Martin, G.R.				
TÍTULO	Osteonectin, a bone-specific protein linking mineral to collagen				
PUBLICACIÓN	Cell 26 (1 Pt 1), 99-105 (1981)				
MEDLINE	82115304				
PUBMED	7934958				
REFERENCIA	2 (bases 1 a 2133)				
AUTORES	Swaroop, A., Hogan, B.L. y Francke, U.				
TÍTULO	Molecular analysis of the cDNA for human SPARC/osteonectin/BM-40: sequence, expression, and localization of the gene to chromosome 5q31-q33				
PUBLICACIÓN	Genomics 2 (1), 37-47 (1988)				
MEDLINE	88256150				
PUBMED	2838412				
REFERENCIA	3 (bases 1 a 2133)				
AUTORES	Lankat-8uttgereit, B., Mann, K., Deutzmann, R., Timpl, R. y Krieg, T.				
TÍTULO	Cloning and complete amino acid sequences of human and murine basement membrane protein BM-40 (SPARC, osteonectin)				
PUBLICACIÓN	FEBS Lett. 236 (2), 352-356 (1988)				
MEDLINE	88312997				
PUBMED	3410046				
REFERENCIA	4 (bases 1 a 2133)				
AUTORES	Le Beau, M.M., Espinosa, R. III, Neuman, W.L., Stock, W., Roulston, D., Larson, R.A., Keinänen, M. y Westbrook, C.A.				
TÍTULO	Cytogenetic and molecular delineation of the smallest commonly deleted region of chromosome 5 in malignant myeloid diseases				
PUBLICACIÓN	Proc. Natl. Acad. Sci. U.S.A. 90 (12), 5484-5488 (1993)				
MEDLINE	93296153				
PUBMED	8516290				
REFERENCIA	5 (bases 1 a 2133)				
AUTORES	Kato, Y., Lewalle, J.M., Baba, Y., Tsukuda, M., Sakai, N., Baba, M., Kobayashi, K., Koshika, S., Nagashima, Y., Frankenne, F., Noel, A., Foidart, J.M. y Hata, R.I.				
TÍTULO	Induction of SPARC by VEGF in human vascular endothelial cells				
PUBLICACIÓN	Biochem. Biophys. Res. Commun. 287 (2), 422-426 (2001)				
MEDLINE	21438978				

Figura 12

PUBMED 1155474JCOMENTARIO REFSEC PROVISIONAL: Este registro no se ha sometido a una revisión final por el NCBI. La referencia de secuencia procede de J03040.1.

CARACTERÍSTICAS Localización/Cualificadores

fuelle 1..2133
 /organismo="Homo sapiens"
 /db_xref="taxon: 9606"
 /cromosoma="5"
 /mapa="5q31.3-q32"

gen 1..2133
 /gen="SPARC"
 /nota="sinónimo: ON"
 /db_xref="locusID: 6678"
 /db_xref="MIM:182120"

CDS 58..969
 /gen="SPARC"
 /nota="Osteonectina (proteína secretada, ácida, rica en cisteína)".
 /codon_inicio=1
 /producto="proteína secretada, ácida, rica en cisteína (osteonectina)"
 /id_proteína="NP_003109.1"
 /db_xref="GI: '4507171"
 /db_xref="LocusID: 6678"
 /db_xref="MIM: 182120"
 /traducción = "MRWIFFLLCLAGRALAAPQOEALPDETEVVEETVAEVTESVGV
 ANFVQVEVGEFDDGAETEEVVAENPCQNHCKHKGKVCLENNTPMCVCCQDPTSCP
 APIGEFEKVCSDNKTFDSSCHEFATKCTLEGTRKGHRLHLYIGPKYIYPPCLDSSEL
 TEFPLMRDHLKAVLVLYERDEDDNLLTEKQKLAVKKINENEKRLKAGDHPVELLAR
 DFEKNYMYIYFVHWQFGQLDQHPIDGYLSHTELAPLRAPLIPMEHCTTREFETCDLQ
 NDKYIALDEWAGCFGIKQKIDKDLVI." SEC ID N°: 1

misc feature 268..339
 /gen="SPARC"
 /nota="FOLN; Región: Dominio similar a N-terminal de folistatina"
 /db_xref="CDD:smart00274"

misc feature 340..504
 /gen="SPARC"
 /nota="kazal: Región: dominio inhibidor de serin proteasa de tipo Kazal. Normalmente indicativo de inhibidores de proteína de serin proteasas. Sin embargo, también se observan dominios de tipo kazal en la parte extracelular de agrinas, que no se sabe que sean inhibidores de la proteasa. Los dominios kazal aparecen a menudo en disposiciones en tandem. Plegamiento alfa+beta pequeño que contiene tres disulfuros. El alineamiento también incluye un solo dominio de transportadores en la familia OATP/PGT"
 /db_xref="CDD:pfam00050"

misc feature 349..504
 /gen="SPARC"
 /nota="KAZAL"; Región: Inhibidores de serin proteasa de tipo Kazal.
 /db_xref="CDD:smart00280"

variación 559
 /gen="SPARC"
 /alelo="G"
 /alelo="C"
 /db_xref="dbSNP: 1053296"

variación 562
 /gen="SPARC"
 /alelo="T"
 /alelo="G"

	/alelo="A"
variación	/db_xref="dbSNP: <u>707157</u> "
	complemento (570)
	/alelo="G"
	/alelo="C"
variación	/db_xref="dbSNP: <u>2304049</u> "
	998
	/gen="SPARC"
	/alelo="G"
	/alelo="C"
variación	/db_xref="dbSNP: <u>1053411</u> "
	1117
	/gen="SPARC"
	/alelo="T"
	/alelo="A"
variación	/db_xref="dbSNP: <u>1060151</u> "
	1451
	/gen="SPARC"
	/alelo="T"
	/alelo="G"
variación	/db_xref="dbSNP: <u>1053660</u> "
	1509
	/gen="SPARC"
	/alelo="T"
	/alelo="G"
variación	/db_xref="dbSNP: <u>3194455</u> "
	1551
	/gen="SPARC"
	/alelo="G"
	/alelo="C"
variación	/db_xref="dbSNP: <u>1054204</u> "
	1655
	/gen="SPARC"
	/alelo="T"
	/alelo="G"
variación	/db_xref="dbSNP: <u>1055210</u> "
	1724
	/gen="SPARC"
	/alelo="T"
	/alelo="C"
variación	/db_xref="dbSNP: <u>1804455</u> "
	1862
	/gen="SPARC"
	/alelo="T"
	/alelo="C"
variación	/db_xref="dbSNP: <u>1804456</u> "
	complemento (1922)
	/alelo="C"
	/alelo="A"
variación	/db_xref="dbSNP: <u>1059279</u> "
	2045
	/gen="SPARC"
	/alelo="C"
	/alelo="A"
variación	/db_xref="dbSNP: <u>1064799</u> "
	2072
	/gen="SPARC"
	/alelo="T"
	/alelo="C"
	/db_xref="dbSNP: <u>1059829</u> "

variación

2104

/gen="SPARC"

/alelo="C"

/alelo="A"

/db_xref=>dbSNP:1804458"

RECuento DE BASES

543 a 533 c 521 g 536 t

ORIGEN

```

1  cgggagagag cgetctgctt gcgcctgccc tgccctgccac tgagggttcc cagcaccatg
61  agggcctgga tcttcttctt cctttgcttg ggcgggaggg ccttggaagc ccttcagcaa
121 gaagccctgo ctgatgagac agaggtggty gaagaaactg tggcagaggt gactgaggta
181 tctgtgggag ctaatcctgt ccaggtggaa gtaggagaat ttgatgatgg tgcagaggaa
241 accgaagagg aggtgggtgg ggaatccccc tgccagaacc accactgcaa acacggcagg
301 gtgtgcgagc tggatgagaa caacaccccc atgtgcgtgt gccaggaccc caccagctgc
361 ccagccccc tggcgagtt tgagaaggty tgcagcaatg acaacaagac cttgactctc
421 tccctgcact tctttgccac aaagtgcac cctggaggga ccaagaaggg ccacaagctc
481 cacctggact acatcgggac ttgcaaatat atccccctt gccctggactc tgactgacc
541 gaattcccc tggcatgccc ggactggctc aagaacgtcc tggctaccct gtatgagagg
601 gatgaggaca acaacctctt gactgagaa cagaagctgc ggggtgaaga gatccatgag
661 aatgagaagc gcctggaggg aggagaccac ccgtggagc tgcctggccc ggacttcgag
721 aagaactata acatgtacat cttcctgtta cactggcagt tggccagct ggaccagctc
781 cccattgacg ggtacctctc ccacaccgag ctggctccac tgcgtgctcc cctcatcccc
841 atggagcatt gcaccaaocg ctktttcgag aactgtgac tggacaatga caagtacatc
901 gccctggatg agtgggcccg ctgcttcggc atcaagcaga aggatataga caaggatctt
961 gtgatctaaa tccactcctt ccacagtacc ggattctctc tttaacocct ccottogtgt
1021 ttcccccatt gtttaaaatg tttggatggg ttgttgttct gccctggagc aaggtgctaa
1081 catagattta agtgaatata ttaacgggtg taaaatgaa aattctaacc caagacatga
1141 cattcttagc tgtaacttaa ctattaagga cttttccaca cgcattaata gtccattttt
1201 tctcttgcca tttgtagatt tgcccattgt cttattggca catgggtgga caccgatctg
1261 ctgggctctg cottaatcac acattgcagc ttcaactttt ctotktatg ttctgtttga
1321 aactaatact taccgagtca gactttgtgt toatttcatt tcagggtott ggctgacctg
1381 gggcttcccc aggtggcctg gaggctggca aagggaagta acagacacac gatgttgtca
1441 aggatgggtt tgggactaga ggctcagtg tgggagagat ccttcagaaa tccaccaacc
1501 agaacgtggg ttgctgagg ctgtaactga gagaagatt ctggggctgt cttatgaaaa
1561 tatagacatt ctacactaag ccagttcat caccatttcc tcttttacct ttcaagtgcag
1621 tttcttttca oattagctg ttggttcaaa cttttgggag caccgactgt cagttctctg
1681 ggaagtggto agogcatct gcagggttct toctctctg tttttggag aaccagggtc
1741 cttctcaggg gctotaggga ctgocaggct gtttcagoca ggaaggccaa aatcaagagt
1801 gagatgtaga aagttgtaaa atagaaaaag tggagttggg gaatcgggtg ttctttctc
1861 acatttggat gattgtcata aggtttttag oatttctct cttttcttca ccttccccct
1921 tgttcttcta ttaatcaaga gaaacttcaa agttaatggg atggtcggat ctacaggct
1981 gagaactcgt tcacotccaa goatttctg aaaaagctgc ttcttattaa tcatacaaac
2041 tctcaccatg atgtgaagag tttaacaaat ctttaaaat aaaaagtaat gacttagaaa
2101 ctcaaaaaaa aaaaaaaaz aaaaaaaaz aaa SEC ID N": 18

```




PubMed Nucleótido Proteína Genoma Estructura PMC Taxonomía OMIM Boo

Buscar para

Límites Vista previa/Índice Historia Portapapeles Detalles

por defecto 20

1: NM_009242. Mus musculus secr...[gi:6678076] Links

LOCUS Sparc • 2072> pb ARNm lineal ROD 07-ENE-2002
 DEFINICIÓN Glucoproteína rica en cisteína ácida secretada de Mus musculus (Sparc), ARNm.
 REFERENCIA NM_009242
 VERSIÓN NM_009242.1 GI:6578076
 PAL. CLAVE .
 FUENTE Mus musculus (ratón doméstico)
 ORGANISMO Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Mus.

REFERENCIA 1 (bases 1 a 2079)
 AUTORES Mason, I.J., Taylor, A., Williams, J.G., Sage, H. y Hogan, B.L.
 TÍTULO Evidence from molecular-cloning that SPARC, a major product of mouse embryo parietal endoderm, is related to an endothelial cell "culture shock" glycoprotein of Mr 43,000
 PUBLICACIÓN EMBOJ. -5 (7), 1465-1472 (1986)
 MEDLINE 86300544
 PUBMED 3755680

REFERENCIA 2 (bases 1 a 2079)
 AUTORES MoVey, J.H., Nomura, S., Kelly, P., Mason, I.J. y Hogan, B.L.
 TÍTULO Characterization of the mouse SPARC/osteonectin gen. Intron/exon organization and an unusual promoter region
 PUBLICACIÓN J. Biol. Chem. 263 (23), 11111-11116 (1988)
 MEDLINE 88298750
 PUBMED 3165375

REFERENCIA 3 (bases 1 a 2079)
 AUTORES Howe, C.C., Overton, G. C., Sawicki, J., Solter, D., Steih, P. y Strickland, S.
 TÍTULO Expression of SPARC/osteonectin transcript in murine embryos and gonads
 PUBLICACIÓN Differentiation 37 (1); 20-25 (1988)
 MEDLINE 88255622
 PUBMED 3384223

REFERENCIA 4 (bases 1-a 2079)
 AUTORES Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. y Hayashizaki, Y.
 TÍTULO Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 PUBLICACIÓN Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCIA 5 (bases 1 a 2079)
 AUTORES Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, K., Akiyama, J., Kishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K. y Togawa, Y.

TÍTULO RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

PUBLICACIÓN Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

COMENTARIO REFSEC PROVISIONAL: Este registro no se ha sometido a una revisión final por el NCBI. La referencia de secuencia procede de XD4017.1.

CARACTERÍSTICAS Localización/Cualificadores

fuelle 1..2079
/ORGANISMO="Mus musculus"
/db_xref="taxon: 10090"
/cromosoma="11"
/map="11 29.9 cM"

gen 1..2079
/gen="Sparc"
/nota="sinónimo": BM-40"
/db_xref="LocusID: 20692"
/db_xref="MGD: 98373"

CDS 90..998
/gen="Sparc"
/nota="osteonectina"
/codon_inicio=1
/producto="glucoproteína rica en cisteína ácida secretada"
/id_proteína>"NP_033268.1"
/db_xref="GI: 6678077"
/db_xref="Locus ID: 20692"
/db_xref="MGD: 98373"
/traducción ="MRAWIFLLCLAGRALAAPQQTVEAEIIVEEETVVEETGVFVGA
NEVQVEMGEFEDGAETVEEVVADNFCQNHCKHKGKVCELDESNTPMCVCQDPTSCPA
PIGEFEKVCSDNKTDFDSSCHFFATKCTLEGTKKGHKLHLDYIGPCKYIAPCLDSELT
EFLMRDNLKNVLVTLIERDEGNLLTEKQKLRVKKIENEKRLKLEAGDHPVELLARD
FEKYNMYIFPVHWQFGQIDQHPIDGYLSHTELAPLRAPLIPMEHCTTRFFETCDLON
DKYIALEENAGCGFIKEQDINKELVI" SEC ID N°: 2

misc feature 297..368
/gen="Sparc"
/nota>"FOLN; Región: dominio similar a N-terminal de folistatina"
/db_xref="CDD: smart00274"

misc feature 369..533
/gen="Sparc"
/nota>"KAZAL; Región: dominio e inhibidor de proteasa S de tipo kazal"
/db_xref="CDD: pfam00050"

misc feature 378..533
/gen="Sparc"
/nota>"KAZAL; Región: inhidores de serin proteasa de tipo kazal"
/db_xref="CDD: smart00280"

misc feature 432..434
/gen="Sparc"
/nota>"señal de glucosilación supuesta (aa 115)"

misc feature 2054..2059
/gen="Sparc"
/nota>"señal de poliA"

RECuento DE BASES 526 a 552 o 518 g 483 t

ORIGEN

```

1 gcattctctg agcccttcag accgccagaa ctcttctgccc gctgcctgc ctgctgcct
61 gctgcctgt gcagagagtt cccagcatca tgagggcctg gatcttcttt ctcccttgcc
121 tggccgggag ggccttgga gccctcagc agactgaagt tgctgaggag atagtggagg
181 aggaaaccgt ggtggaggag acaggggtac ctgtgggtgc caaccagtc caggtggaaa
241 tgggagaatt tgaggacggt gcagaggaaa cggtcgagga ggtggtggct gacaaccct

```

```

301 gccagaacca tcattgcaaa oatggcaagg tgtgtgagct ggaogagagc aacacccccca
361 tgtgtgtgtg ccaggacccc accagctgoc ctgctcccat tggcgagttt gagaagggtat
421 gcagcaatga caacaagacc ttcgactott cctgccactt ctttgccacc aagtgcacccc
481 tggagggcac caagaagggc cacaagctcc acctggacta catcggacca tgcaaataca
541 tcgcccctg cctggattcc gagctgaccg aattccctct gcgcatgcgt gactgggtca
601 aaaatgtcct ggtcaccttg tacgagagag atgagggcaa caacctctc actgagaagc
661 agaagctgcg tgtgaagaag atccatgaga atgagaagcg cctggaggct ggagaccacc
721 ccgtggagct gttggccga gactttgaga agaactaca tatgtacatc ttccctgtcc
781 actggcagtt tggccagctg gatcagcacc ctattgatgg gtacctgtcc cacactgagc
841 tggccccact gcgtgtctcc ctcatcccca tggaaattg caccacacgt ttctttgaga
901 cctgtgcct agacaacgac aagtacattg ccctggagga atgggcccgc tgctttggca
961 tcaaggagca ggacatcaac aaggatctgg tgatctaagt tcacgcctcc tgcctgcagtc
1021 ctgaactctc tccctctgat gtgtcaccoc tccattacc ccttgttta aaatgtttgg
1081 atggttggct gttccgctg gggataaggc gctaacatag atttaactga atacattaac
1141 ggtgctaaaa aaaaaaaaaa aacaaagtaa gaaagaaact agaaccacag tcacagcatt
1201 ttcccacata actctgaggg catggcccat ccacagocct ctggtccct gcactaccca
1261 gtgtctcact ggctgtgttg gaaacggagt tgcataagct caccgtocac aagcacgaga
1321 tatctctago ttctatttca attttgcatt tgactcttaa cactcaccac gactotgtgc
1381 ttatttcatt ttgggggatg tgggcttttt cccctgggtg tttggagtta ggcagagggga
1441 agttacagac acaggtacaa aatttgggtg aagatactgt gagacctgag gacccaccag
1501 tcagaaccca catggcaagt cttagtagcc taggtcaagg aaagacagaa taatccagag
1561 ctgtggcaca catgacagac tcdacagcgc ccgggacctt gctgtctctc cgactottcg
1621 ggcttttct tccatgtttg gctgttgggt ttagtttttg tgagccatgg gtgggocaga
1681 acatcactca actgcaattg ggctttcagg ttcttgccgg gagctctagg cactgggagg
1741 ctgtttcagg aaagtgagac tcaagaggaa gacagaaaag gttgtaacgt agaggaagtg
1801 agactggtga attggtttga ttttttccac atctagatgg ctgtcataaa gtttctagca
1861 tgttcccct cactctccc caccocctgc cacttgaaac cttctactaa tcaagagaaa
1921 ctccaagcc aacggaatgg tcagatctca caggctgaga aattgttccc ctccaagcat
1981 ttcattgaaa agctgcttct cattaaccat gcaaacctct acagcgatgt gaagagcttg
2041 acaagtcttt caaataaaa agtaacaact tagaaacgg

```

SEC ID N°: 19




PubMed Nucleótido Proteína Genoma Estructura PMC Taxonomía OMIM Boo

Buscar para

Limites Vista previa/índice Historia Portapapeles Detalles

1: AB088113. Homo sapiens SC1,...[gi:27544421] [Links](#)

LOCUS AB088113 14468 pb ADN lineal PRI 08-ENE-2003
 DEFINICIÓN SCI de Homo sapiens, genes de OTF3 para factor de transcripción .19, factor de transcripción 1 de clase 5 de dominio POU, cds completo.
 REFERENCIA AB088113
 VERSIÓN AB088113.1 GI:27544421.
 PAL. CLAVE .
 FUENTE Homo sapiens (humano)
 ORGANISMO Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCIA 1
 AUTORES Shiina,T., Ota,M., Katsuyama,Y., Hashimoto,N. e Inoko.H.
 TÍTULO Genome diversity in HL: A new strategy for detection of gentic polymorphisms in expressed gens within the HLA class III and class I regions
 PUBLICACIÓN No publicado
 REFERENCIA 2 (bases 1 a 14468)
 AUTORES Shiina,T.,
 TÍTULO Remisión directa
 PUBLICACIÓN Remitido (08-JUL-2002) Submitted (P8-JDI-2002) Takashi Shiina, Tokai University School of Medicine, Molecular Life Sciense 2; Bohseidai, Isehara, Kanagawa 259-1193, Japón (E-mail:tshiina@is.icc.u-tokai.ac.jp Tel:81-463-93-1121, Fax 81-463-94-8884)
 CARACTERÍSTICAS
 fuente Localización/Cualificadores
 1..14468
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cromosoma="6"
 /mapa="6p21.31"
 /línea celular="LXT3"
 /tipo celular="célula B"
exón 1066..1680
 /número=1
gen unión(1443..1680,3384..3942,4414..4654)
 /gen="SC1"
CDS unión(1443..1680,3384..3942,4414..4654)
 /gen="SC1"
 /nota="proteína de unión a AND necesaria para la progression al cliclo celular tardío"
 /codon_inicio=1
 /producto="factor de transcripción 19"
 /id_proteína="BAC54945.1"
 /db_xref="GI:27544422"
 /traducción ="MLPCTQLLRIGGGGGGLYTFHPAGAGCTYRLGHRADLCVVAL
 RPQQEPLISGIHAELBAEPGRGDDWRVSLDRSSQGTLVNNVRLPRGHRLELSDGDL
 TFGPEGPPGTSPEFYFMFQVRVKPDFAITIPRSRGEARVAGFRMLPSQGAEP
 RPLSTFSAPKATLLNSIGSLSKLRPQSLTFSPSWGGPKSLPVPAPPGEVGTTPSAP

PQRNRKRSVHRVLAELDDESEPPENFPVLMFPRKRLRVDKAPLTPTGNRRGRPKYF
 VSAPMAPPVGGGEPCAAFCCLPQEBETVAVWQCDCDVWFHVACVGCSTQAAREADF
 RCPGCRAGIQT" SEC ID N°: 3
 3384..3942
 /gen="SC1"
 /numero=2
 4414..5723
 /numero=3
 6283..6805
 /numero=4
 gen complemento (unión(6539..6305,7068..7226,7511..7641,
 7875..8115))
 /gen="OTF3"
 cgs complemento (unión(6539..6805,7068..7226,7511..7641,
 7875..8115))
 /gen="OTF3"
 /nota="proteína de unión a AND que contiene homeodominio
 de tipo POU"
 /codon_inicio=1
 /producto="dominio POU, clase 5, factor de transcripción
 1"
 /id_proteína="BAC54946.1"
 /db_xref="GI:27544423"
 /traducción="MHFYRLFLGATRRFLNPNKGEIDNWCVVYVLTSLLEFKIQSQDI
 KALQKLELOFAKLLKOKRITLGYTQADVGLTLGVLFKGVFSQTTICRFEALQLSFKNM
 CKLAPLLQKWEADNENLOEICKAETLVQARKKRTSIENRVGNLENLFLQCPKP
 TLQQLSHIAQQLGLEKDVVRVWFCNRRQKGRSSDYAQRDFEAAGSPFSGGPFVSFP
 LARGPHFGTFGYGSPRFTALYSSVPFEGEAFPPVSVTTLGSPMRSN" SEC ID N°: 35
 7068..7226
 exón número=3
 exón 7511..7641
 exón /numero=2
 exón 7875..8216
 exón /numero=1
 RECUENTO DE BASES 3555 a 3859 c 3746 g 3308 t
 ORIGEN

1 tctaaggctg ttctgatctc ttcattctgc ccttcacctg gccctgtac ccccttcacc
 61 tttggaccac ttgaaccctc ccaggaccac cgtccagacc ctcccgccc ccaaccgact
 121 cttcccgacc gtcccttacc aaccggagga gccccctact gcgctttggc cacaccctct
 181 agcctcgcct ccgggcccgc cctctgccc tgaccgcgc tgcgcaaggc gggcgcccta
 241 aagtcctatt tcactctgtt gggaggaggg ggaagggtgt acgcaggctc agtggcgtct
 301 aaatttgggc caactaatg cgtcggagca tctccggccc caggcggtct ctcctcactg
 361 cggcaaccgc ggaactctg tgaactaatc agaaaaagtg gaaggcgagg gatcttgggg
 421 cgtcttccca tggcgaggaa gagaacaaat gagctggcca atcggaacgc gacggggggc
 481 gggctcgcct ggcgggaagt tcgggcccgc gaattccgaa ggaggggtag gcgctgccgc
 541 cggcgagagg ccgcgccct cctggcccgc gcttcttggc tgtcaaacag atgcagcaac
 601 gtcggctcct ggcgaggagc ccaagggtct cccggatccg ccgcacaggg tggcactgct
 661 tgaagaggag gctactcgga gactgcgcgc cgggggtaga tccgaaacgc ggcctggggc
 721 gagtgggaaa aggcggggtg tgccttgcct gatcgccggg agctccttcc tgtttttatc
 781 ccaactagag aagccgggaa gtagggtttt aggtccaat tgttggagta ctttaaggact
 841 cgtttgcact ttcttttggg ggtgacagt ggttgcatt cctcggagg ttcaaactagt
 901 tatgagttag ggtttggcca gaagatggg ggcagggcaa gcaggagtgc tctattagga
 961 taagcaagt ttgacaggaag aagctgctct tctccgaatt acacagagg gatgtgtctg
 1021 tattgcaagt agactgtgt ataacaggac ctcttccccc cgcgccccc accccgaac
 1081 acacaggagc tgcctaaagt atccttgcct tgcagattgg aggtcccaa aatattttgt
 1141 gatctgagga tccagotcaa gtgagggtgc ataggagctg ttcctgagtt tgcattgcac
 1201 ggagacotct ctggaaattt tcatttgcga gtcggcttaa ccaattttgc attgagtcct
 1261 aggttgcttg cactctgaat ttgggctatt caggtagtgt gctcaagtgt gaaaccgcac
 1321 acagcacaac tcaagtttgc atcagactgg gaagcgaact taagocagag gtgcttggcc
 1381 caggagtggg aaaggaaatg gatgcctgaa gtggaaggag tgggtgcagag ggggcaccgc
 1441 ccatgctgcc ctgcttccaa ctgctgcgca tagggggcgg caggggcggt gatctctaca

1501 ccttcacccc ccccgccggg gctggctgca ootatogett gggccacagg gccgacctgt
 1561 gtgatgtggc cctgcggccc cagcaggagc ctggcctcat ctctgggac caccgccaac
 1621 tgcattgccga gccccgggggt gatgactgga gggtcagcct ggaagaccac agcagccaag
 1681 gtgagcatta agcagggtag ctttgccctt ggggtggtga agcgccaggc tggatagagt
 1741 aaggtotcca caagacctg ctgctgctt cccatctctc catcagattg gatggatagt
 1801 cgtggtccag accttcattt tccaccaga agtgtgcaca gtcagaaagt ctctgccaga
 1861 ctgacctttt ttggctccgt ttagctcata caggacctgg gatatacaca gaaagatata
 1921 acagtgggga tgttctgagg coactagagg ccaagtttag aottgattca gtttcacagt
 1981 ttgctgaggc actctgttcc tgggttaggg cagttctatg ttgaataatg tttttaataa
 2041 tctgggcatg tctttctccg tgaactgagg cagtttagct cagaaagcct agattcacat
 2101 ttgagttttg ccaactgctc ttggtaaagt cagctgtagg agtggtatgg ttatttagact
 2161 atagttagca acatttcattt agtgcctact gttatgagcc agggccctatt ttaagtgtat
 2221 tgaatgtagg tggtaactaa attatctcca tttacagtaa aggaaatga ggcacaaaaga
 2281 ggttaaggaa cttgtccagg gctgggcatg gtggtttaca octataatcc agcaactttg
 2341 gaggctaagg cagggtggat cacttgagct caggagtctg agaccagcct gggcaacatg
 2401 gtgaaaacct gtctctacca aaaaattaat taatttttta aaaaagoot gggcggggtg
 2461 gctcagcctt gtaatccacg cactttggga ggcagagatg ggcagatcac gaggtcagga
 2521 gttcgagacc atcctgacca acatgttgaa acccatctg tgetgaaaa aaaatacaaa
 2581 aattcagccg gtgtggtggc gtgcacctgt aacccagct actcaggagg ctgaagcagc
 2641 agaatacatt gaacccggga ggcggagggt gcagttagct gagatcgcac cactgcactc
 2701 cagcttgggc gacagagoga gactccatct caaacaaaca aacaaacca aagcttgccc
 2761 agggctacat aactggttag tggtagagct aggatctgaa cagactggag ctgggggaga
 2821 gtgagcatgt ttgaaaactg gacottaggg cggggcaogg tggctcacgc ctgtaatccc
 2881 agcacttttg gaggctgagg cgggcagatc aggaggtcag gagtatgaga ccagcctggc
 2941 caaataggta aaacctgtc tctgctaaaa ataaaaaat tagccagacg tgggtggaca
 3001 tgcctgtaat cccagctact caggaggctg aggcaggaga attgcttgaa cctgggaggc
 3061 ggttcaagct tgggcaatag agcaaaaact catctcaaaa aaaaaaaaag aaagaaaaaa
 3121 aaagaagaaa gaaagaaaat tggaccttag gacagtggag gcagggatcc tttgtaggaa
 3181 agcaaaagaa acacagactt gttccttagct gacaaggagt gtaactgctg gtaacctgtc
 3241 cctgctgagg ggttaggat gtgaggaga atctgaatao agtttcatat tcttccocag
 3301 aaatcataca gatttctcca ctctgactc tggctatttc tgtttttgtc ctccatattt
 3361 gctggtgccc ccacatcaa caggtacttt ggtcaataat gtcgactcc caagaggtca
 3421 caggctggaa ttgagttagt gaggacctct gacctttggc cctgaagggc cccaggaac
 3481 cagccctctg gatttctact tcatgttcca acagtagcga gtcagcctc aggaactttg
 3541 tgocattacc atcccaoggt ctaggggaga agcccggtt ggggctggt tccggcctat
 3601 gctgcccctc cagggggctc cccagggccc totcagcacc ttctcccctg ccccaagga
 3661 cactctgac ctaactcca taggcagctc cagcaagctc cggccccagc cctcaccctt
 3721 ctcccctagt tggggtggac caaagagcct gctgttccc gcccacctg gggaggtggg
 3781 gaaacagcct tctgctccac cccaaagcaa tcgggggaaa totgttoacc gagtgttggc
 3841 ggaactggat gatgagagt agcctctga gaacccgcca ccggctctta tggagccag
 3901 gaagaaactc cgtgtagaca aagcccaact gactccact gggtaagtgg agtctcact
 3961 tggccctctc agtgttttao tgccttttga ttccttgtat cccataggctg tgaggaggto
 4021 cccctgctg gggggatggg caccggaggt ggaatagatg gaatggcaag acctgggtta
 4081 gctctgatag gaaaagaaaa atatgtgacg gagaacatga gaggtgggt ggggcagtgc
 4141 ttataaaaaa accggagtga gcatgtctg ctttttacat tcatatggct ttaaccccat
 4201 tottctagt cctaaggatg gggaaacttt aggtccacac fagaggtttt tagggccacc
 4261 ctatgtgttt ttaaggacag agtccaggct caocttagtt ctcagaccac tgtgctctg
 4321 tggcctcaco ctatgaocag ccatagggtg gcaaggctca ggcctctcc tacaggtttc
 4381 cggtagacct tgtgtctgtg tcacttccct cagaaatcga cgtggccgtc ctgggaagta
 4441 cccagttagc gctcccattg ctcccctgc agttgggggc ggggagcct gtcagctcc
 4501 ttgttgcctc ctgcccaggg aagagacagt ggcctgggtt cagtgtgat gctgtgact
 4561 ctggttccat gtggcctgtg ttggctgcag catccaggct gccagggagg ccgactccg
 4621 atgcccaggg tggcgggctg gcatteagag ctaaggctca ccgccaagg accatcggac
 4681 acacctgccc atgagtagac acagcagoga gcaaatagggt ctgataaata ccccccctcc
 4741 cttccctccc caagaggga tgaatcagag gaaagaggat ggattgatgt ggactcatte
 4801 agggcctgga gcagacctg gtggccaaga cagaagagat ggtttctgc caaagatatt
 4861 ggcacotcca ggaattggc agttagctgg aagtccac tattacaaga catagggca
 4921 tgttgcctg gacaccaga tatctgtagt cagagcaoot atcagttgca aaagccatgc
 4981 ctgcaaccga tggaaaatgt aagaggagt tottaagggt cttggtggca tcaccaagg
 5041 cattctggga aaacctaggg cctggcccca aaacttccct actctgtggo tagtccctgt

5101 gccacacaaa tggtagcgac ctggccttttc acagcctttgc ttttatttcc aagtcacagga
 5161 caagccgctt cattcactcc tgggcattta ctottcttgt gggctctgtga tattccttgc
 5221 ttccagggga gaatgtgctt ggcaaggctc ggagaactaa ttcagaatct taggggaagg
 5281 ggagagatgg aaatacaaac ctgcttactg gaaaggtgca aatataatggg ttgagctgga
 5341 ggttaggaata caggttaetta aggtttctag ttttaaggga aacagatcta ttgccattta
 5401 aataaggtaa ctgggatttg gtttaagttca caaagatagc agaagattta tttacaggct
 5461 tcacctgtac tgtcagggca agagaaagcc tggtaaacca gctacagcag tttaccagtg
 5521 tgatggctgt gacacagctc cactccacgg gtggacacag cagagggcaa ctgggctggc
 5581 ctgggttcagt gtgaatcaaa ccgcttaacc caccacatgg acatgtgatt ttcttttgtg
 5641 agccttacac caagccaaa c tattgtcaaa gcatcatttc tatagaaata aagccttacc
 5701 ttgacctgtt ctattaaaac ctgcccacac cgcctttcc taactagatt taatgagccc
 5761 aagttttttt acatggaaga aatgactctg gggcaaaagc cctaatgaa ctagtggcag
 5821 agccaggaat aaaacttgag taactaatga gtcacttatg ggcagagtat gcaaaaacct
 5881 taagtggaaa ccaatagac cctggtatca agaaagcaca aagtattaat agaagtttct
 5941 ggttgggggtg atctaggttc aacagaaata agatgatttc taagtataaa gccatttaag
 6001 aattccagag taggggtggga aagcaaaaag ccagctctga acaggtaaca gctacatggg
 6061 gactgagctc atgggcaaaa gttcttgcac cacaggcttt tgggaaactag cctatccag
 6121 ggccctgtac aaataaactt ggctgcaatc ccagctctcc ctctgatgtt gttgacctt
 6181 aaggagtgtg aatggcacct tagtttcagg gtcacttggg tatgagcatt ggtatttccc
 6241 atccccaact cagtaactga aggacaaaac aagataagtg tgtctatcta ctgtgtccca
 6301 agcttctttt ttaagaaaa aagtgtatca tgatgtggga ttaaaatcaa gagcatcatt
 6361 gaacttcacc tccctccaa ccagttgccc caaactcccc tgccccacc ctttgtgttc
 6421 ccaattcctt cttagttaa tgaagaaact aatcccaaaa accctggcac aaactccagg
 6481 tttttcttcc ctagtctctc cctcccccct gtcccccatt cctagaaggg caggcacctc
 6541 aytttgaatg catgggagag cccagagtgg tgacagagac agggggaaag gcttccccct
 6601 cagggaaggg gacogaggag tacagtgcag tgaagtggag gctcccatag cctggggtag
 6661 caaatgggg cctggggccc agaggaaagg acactggctc cctgagaaa gtagacccag
 6721 cagcctcaaa atcctctcgt tgtgcatagt cgetgcttga tgccttgccc ttctggggoc
 6781 ggttacagaa cccactcgg accacctgcc agtgaatgac agaaaggaga atgacattag
 6841 acaatgagct gagagaaggg cctgactctg cttggacatt ctatccaaag ccaacagccc
 6901 tagagcagtt agaggaggac attagagaat gagctgagac aggcctgact gcttggacat
 6961 tctgtcccaa gccaaagccc cttagagcagt tggaggagcc agagctaggg aaagcgaggt
 7021 ggtgacaggg gaaagagatg gagcccgccg agagacatgg cactcacatc cttctcgagc
 7081 ccaagctgct ggccgatgtg gctgatctgc tgcagtgtgg gtttcgggca ctgcaggaa
 7141 aaattctcca ggttgctctc cactcgggtc togatactgg ttogctttct ctttcggggc
 7201 tgcacagagg tttctgcttt gcatatctgt gcaggtggga aggggggtgac aaggggcaagc
 7261 tttggacttg ctgagtaaca gcatcacagg ggtctgtgac tagatgtgtc agcagagcca
 7321 ggtggtggtg tgaaaaggca ggtccttggg aggggtggct ctgggaactta tccagcaga
 7381 actgaggaat ttcactccat cccactgaga accactgcac caaagacgga gactacag
 7441 ccagtgtagg aagcaatgga aattaggoca agaaagggaa ggtccccggg tatccccctc
 7501 ccacccttac ctccgaaga ttttcattgt tgtcagcttc ctccaccac ttctgcagga
 7561 agggccgag cttaacacatg ttcttgaaag taagctgcag agcctcaag cggcagatgg
 7621 tggtttggct gaatacctc cctgggggag gccagtcaaa agageagcaa aatgagggag
 7681 cagcagggc cttgtgacc ctgagatcca agcttaaaa ctcttccag agggagctca
 7741 aagoccaaag atottctccc tctccctact cctcttcatg ggtgagggta gagtgtgcc
 7801 ctgcccctcc ccactaggtt cagggatact ccttagaggg gagatgggt cagaatctgc
 7861 agagggggaa ccaccaaata gaacccocag ggtgagcccc acatcggcct gtgtatatcc
 7921 cagggtgato ctcttctgct tcaggagctt ggcaaatgac togagtctt tctgcagaga
 7981 tttgatgtcc tgggaactgga ttttaaaagg cagaagactt gtaagaacat aaacacacca
 8041 gttatcaato tcccctttcc attcgggatt caagaacctc cgtgtggccc caaggaaatag
 8101 tctgtagaag tgcactgccc ttccaagctg cccacctaac ttctagaaat aaactacca
 8161 caaatgtcat tcacccattc cctgttcact gactcatgca tgaacaaaag gactactctt
 8221 cccccagaaa ctggcacatc caagggatgc agagcatggg gaaaggacag aaagagagac
 8281 cctggcctcg aggagaacac ctgtcaggtt atgaaggta gaagttcttt gctgggocg
 8341 gtggctcatg cctataatto cagcactttg ggaggccgag gtgggcagat caccaggtca
 8401 ggagttaag accagcatgg ccaacatggt gaaacccogt ctctactaaa aacacaaaaa
 8461 ttagctgggc acggtggcac gcaoctgtaa tcccagctac tcaggaggct gaggcaggag
 8521 aatocactga accggggagg cggaggttgc agtgagctga gatcacgcca ctgcactcca
 8581 gcctgggtga cagagcaaga ctctgtctca aagaaaaaaa aaagaaagat agttcattta
 8641 atacctgcaa aattctctca ctcaagtatc accccagtt taaggatgtt ttgagattag

8701 agaaatagat aagctgctaa gttctgggtt aattaaaaag gaagagcatc atgtctcaga
 8761 agctaaattc agtatatact ctcccagct tgccttgagg gtcccacaaa ctataacatg
 8821 gcatgcatac acacaaacac agcaaaaaag taacagggtg cataagaatg gataaagtgc
 8881 tttgtgtgta cttactcctc atttttttaa ttgattatcc ctcatcttta ctgtatcttt
 8941 ttcactatag aggcaccta attgattttt aaattcaga gatttatoga gcaccttcta
 9001 taagccagcg gctatacaaa gtggacaaag agccctgaca tcagcatga cagaagtgtt
 9061 attoggcact tgttcttoaa gttgccact tggatctott ccaagtgoac tttccttttt
 9121 tccctgcoct ataacttttt aataataaa ttccactcct gctctgaaaa ataaaaagt
 9181 aaataaaata aaaaatggcg aggcacagt gctcatgtot gtaaatocta gcactttggg
 9241 aggcacaggt gggcagactg cttgagccca agagttagaa agcagcctgg gtaacatagt
 9301 gagaccggtg ccgccccttc tcccacccct gctgctcta tttaaaaaat atatatatat
 9361 tatggaaaaa agcaaacag tccgggcgca gtgctcatgc ctgtaatccc ttcactttgg
 9421 gagggcaagg tgggtagatc acttgaggto aggagttcaa gactagcctg gtcaacatag
 9481 tgagactctg tctctactaa aaatacaaaa attagctggg catcatggcg ctcccctata
 9541 atcccagcta otcaggagcg tggggcagga gaattgottg aacotaggag gtggagtgtg
 9601 cagtgaacca agatgcaccc actgcactcc agcctgaggg acagagttag actccatctc
 9661 aaaaattaaa aaaaataaa agcagcttat aggagtaggg taaaggaggg aaggagattt
 9721 tggaggaggg tgacactttt aaagacagag aagggtattg tttgagcaaa ggacaagagt
 9781 ctaatgtggc aaggccctga agtgggcctt ccagagccca aagctggtct ggtggttagg
 9841 tagatctgtg tgacagcata gtgactttgt tttagtccaa gtgaaatgat ctctcacccct
 9901 ttttctcccc ccccaagagc gaatctcgtt ctatogccca ggctggagtg ctgtggcgtg
 9961 atcttggcto actgcaatct ccgccttctg ggttcaagct attctgctc agccgctga
 10021 gtagctggga ctacaggcac ccaaccacat gccoggotaa tttttgtatt tttagttagt
 10081 atgggggttc accatgttgg ccaggctggg daggagacot caagtgatct gtccaccttg
 10141 gottcccaaa gtgctgggat tacaggtgtg aaccacccga cctagccctc cctttttttt
 10201 tttttttttg agagtttogc ttttgttgcc taggotggag tgcactggcg cgtctcggc
 10261 ctacogcaac ctacatctcc caggttcaag cgattctcct gctcagctt cctgtagtgc
 10321 tgagattaca ggcacgtgto accaogccca gctaattttg tatttttagt agagtgggg
 10381 tttogccatg ttggtcaggg tggactogaa ctcccaacct cagggtatc gctgctctog
 10441 gctcccaaaa gtgctgggco acacctttta aaacaotgac tctagttgac gtgttggcca
 10501 cagacagtag ggaggaaagca gtataatttg agaagctact ggggtaatcc cagcagagat
 10561 gatgttggct gaggccaggg ttaggttgtg attgattcag gatgtttctt aaggatagg
 10621 tgtaggacgt gaaagaaact gaggatgact gggtttggco ttgagcaact ggtgtatcag
 10681 ggtggagcag ttoagggagc catcacaaga gacagaaaac ggggtagtca tctgtgtct
 10741 aaatggcatt taagccttga ggggtgggtg gaggagaggaa ggttagatag agcagaggtt
 10801 gaaggactga ggcctggggc atgocatatg aggtgocgg cggacagagg tgcacagcta
 10861 gtgagaaaaa aacaaggcct tttttagtgc ctgaagcctc aaggaaagtgt tccaatgggtg
 10921 cttgatcata tcaatttcaa ataggctgtt ttcatcccca acttctgtct agccaataac
 10981 tcaaaactgat aaatgcccctc tgotatctct gattttccaa attctgtttt ggggtttttg
 11041 aataaacact ggtccaaatc ctgccttcat catttagcag ttaaaaoccg ttaaatagg
 11101 taataatacc toccoctagg agattttgtg ctggttaatg agataatgat gtataaacgg
 11161 agcacacagc caggcactta ggaagtggac cacaattgoc agccattatc attcaaggct
 11221 cagcagtgac ctocctggaa gaggttgggg cttctcggtc actccagada ccagtcaao
 11281 cttctctgtg ggtctcaagg cttagtattt aatctctaat tgcctacaot tgtcgccttg
 11341 gaggactgga agatacatct ttaatagtcc tcaagcaggc tggatgcctt caatcccgca
 11401 gcagctctat atttgcaaat ggcctggaga aatctctcac catttttctt gtttaoaact
 11461 ttggaactga ggtgaaatc aatcaaaatc cagctttcta caagggtgtc cagggtgtgc
 11521 acctaaacac agtggccagt cattggcctg aggcagagat ccgggggaag caagccctat
 11581 acttgactgg aggtaaaccc agctcacaac gcgcacacac acagcccaaa caggagatcc
 11641 tatcagaaac gagtacacac ctagacttto aggaacaata atcctggaat gaggactgtt
 11701 tttacccctc ggetatgott aacccaaagg ccaaaatctt gggcttgata aggggtcaat
 11761 tttcaagcag gactaagggt gggaaagggg gctcaaaacc aocccaagct gggctgtgtg
 11821 ctgggocagt aatgagtga cagacccctg gcaggcctag gagatgtgag agacccctgac
 11881 aagggtggg ccagacagag caaaggccag cctgggcccag cttccgactc tcccaggcog
 11941 ctctgcccctc acctgcagtt gtctcttcga aatccagott ccagttccca cctggcccct
 12001 gctgcccag gctgctgca gttgatacac aocccctcct ggcaggggca gctgacccct
 12061 cctgctcctc tctgggttgc caggctgtgg cagctgcagg tgacccactc cccatcaggc
 12121 gctgctgcta tgaacaccc cccacacccc aaccccgctg aagctcaact gctcctccg
 12181 ggttttgcctc cagctctctc tctctcagct tcacggcacc aggggtgacg gtgacgggct
 12241 ccggggaggc cccatcggag ttgctctcca ccccgactcc tgccttggcc tcaggctcag

12301 aggtctccaa gccgcottgg ggcaactagcc coactccaac ctggggccca cagtaogcca
 12361 tccccccaca gaactcatat gccggggggc atgggggaat ccccccaccc tcagagcctg
 12421 gcccaccccc cggcccgatt cctggccctc caggagggcc ttggaagctt agccaggtec
 12481 gaggatcaac ccagcccggc tccggccccc ctggcccatc acctccaaca cctggagggg
 12541 gcgagaaggc aaaatctgaa gccagggtgc ccgccatggg gaaggaaagg gccccaagcc
 12601 gggggccctgg tgaatgagg gcttgogaag ggactactca acccctctct cctcccccag
 12661 tcccaccacac tagccttgac ctotggcccc gcccctgga tgggtggagg agagggaggt
 12721 ggggggagaa actgaggcga aggatgtttg cctaattgtg gtggcaatgg tgtctgtgga
 12781 aggggaaaac cgggagacac aactggcgcc cctccaggac ctcagtgcag gtccccaca
 12841 gaaacttttt ttatttttat tttttaagac aggtctcac tttgtggcc agactggagt
 12901 gcagtggagt acaatgatgg ctcaatgtag cctcgatcta ctgggcaaaa gcaatccttc
 12961 tgctccagcc tcctaagtgg ctgggactac aggtttggac cactgtgccc tgttagtttt
 13021 tttattttta gtagagatgg ggcccttgcta tgttaccag gctgtcttg aattcctgtc
 13081 ctcaagaaat cctcccgct ctgcgcgcca gtgtcatgat taaaggcgtg agccaccaca
 13141 cccaacttte aactcccaac ccgctccctg gcaactctctc aggtctgca catcccagct
 13201 gtctggaatc actccacac ctccatgttc ttcaggaaac caggtgttg accccctctc
 13261 cacagacetc tggcaactgt ccttcagggg ccagtcaccc tctcagctcc tcaaatttat
 13321 tgaatgtgtg tgtggcgcta tccctcaatg catcaacagc cataagcaca atggccagct
 13381 gctcccttat gcttccccc gatccatcca gaatcctagg cattcccatc ccgatactgg
 13441 ccaaatccag ccaccccga gctgggtgc ctggcaccat ctgccagcc tgcctaattt
 13501 caccocatct tcaagagtag actgccagac aaggcctccg tgctatatcc cccacacccc
 13561 cateccccc cccctccgtc ttccagaatc agactccaga ctctcctcat ctaacagact
 13621 aaggggttgg cccctaottc cccttcaagg gaccagactt tggactgatt gggcctcagt
 13681 ttcccaacct ttgtgaaac agagtataaa gacacccgt ttgggcccc tccactatgg
 13741 aacctgcaca tcaggttcct tgcctccctc tcaaccaaaa ctccagacatc taataccacg
 13801 gtaggcccg ttctccctcc cccacctccc tggcccaggc ctccagccct agggcctggg
 13861 tggggaaaac caggggttgg ggggtgtgga gaaaaaatat ctgacttcag gttcaaagaa
 13921 gcttgggagg gactggggga agggggcagg acaatggcct tggctggaca atcccggtcc
 13981 ccagaggggg cagctctaac cctaaacaag tgcctcaacc ttgaatgggc ctggatggct
 14041 cccctgggga ctgcttctg ctccccaacc cccagtcctc aatccctca cacagaatcc
 14101 ccttcagaga cgtctaaaagg agctccagca accccctct gcaatccct caaagactga
 14161 gcttcagagc ggcaccaagg gccccctaca gggacctagg tatctagttc ctccctctc
 14221 tgggggactc aggcgtccag ctteatctgt catccctccc cgagcccgga agattgaggg
 14281 atgtgctttg tttagtggg ctggctggca gaaagacgca gaggaggtgg cagtgattt
 14341 gtggaggggt gcagggaagg tgccttaagc tccccttcag ggtctgttt tctgggcctg
 14401 gctgagtat cctgaggtc atgtgtctg tctagtgtt gattctgtt gcaagagaat
 14461 agccaaog SEC ID N°: 20




PubMed Nucleótido Proteína Genoma Estructura PMC Taxonomía OMIM Boo

Buscar para Ir

Limites Vista previa/índice Historia Portapapeles Detalles

por defecto 20

1: BC033086. Homo sapiens, Sim...[gi:21620013] [Links](#)

LOCUS	BC033086	3Q21 pb	ARNm	lineal	PRI 27-JUN-2002
DEFINICIÓN	Homo sapiens, similar a factor de transcripción 19 (SC1) ; Similar a factor de transcripción #, clon MGC:45652 IMAGE:3160434 , ARNm cds completo				
REFERENCIA	BC033086				
VERSIÓN	BC033086.1 GI:21620013				
PAL. CLAVE	MGC.				
FUENTE	Homo sapiens (humano)				
ORGANISMO	Homo sapiens				
REFERENCIA	Eukaryota; Metazoa; Chordata/ Craniata; Vertebrata; Euteleostomi;				
AUTORES	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TÍTULO	1 (bases 1 a 3021)				
PUBLICACIÓN	Strausberg, R.				
	Remisión directa				
	Remitido (25-JUN-2002) National Institutes of Health, Mammalian				
	Gen Collection (MGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
	USA				
OBSERVACIÓN	URL del proyecto NIH-MGC: http://mgc.nci.nih.gov				
COMENTARIO.	Contacto: servicio de ayuda de MGC Email: cgapbs-r@mail.nih.gov Obtención de tejido: ATCC Preparación de biblioteca de ADNc: Rubin Laboratory Biblioteca de ADNc organizada por: El consorcio I.M.A.G.E. (LLNL) Secuenciación de AND por: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.be.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Seott Zuyderduyn, Marco Marra. Distribution de clon: La información de distribución del clon de MGC puede encontrarse a través del consorcio I.M.A.G.E. /LLNL en: http://image.llnl.gov Serie: Placa IRAL: 43 Fila: a Columna: 8 Se seleccionó este clon para secuenciación ed longitud complete debido a que pasó los siguientes criterios de selección: ARNm coincidente gi: 6005891.				
CARACTERÍSTICAS	Localización/Cualificadores				
fuelle	1..3021				
	/organism="Homo sapiens"				
	/db_xref="taxon.9606"				
	/clon="MGC:45652 IMAGE:3160434"				
	/tipo_tisular="cerebro,neuroblastoma"				

```

/ lib_clon="NIH_MSC_19"
/ laboratorio_anfitrión="DH10B-R"
/ nota="Vector:pOTB7"
CDS
422..1459
/ codon_inicio=1
/ producto="Similar a factor de transcripción 19 (SC1),
factor de transcripción similar #"
/ id_proteína="AAH330B6.1"
/ db_xref="GI:21620014"
/ traducción ="MLPCFQLLRIGGGRGGDLTYFHPFAGAGCTYRLGHRADLCDVAL
RPQOEPGLISGIHAEHAEPRGDDHRSLEDHSSQGLVNNVRLPRGHRLELSDGDL
TFGPEGPGTSPSEFYFMFQQVRVKQDFAAITIPRSRGEARVAGAGREMLPSQGAPO
RPLSTFSPAPKATLILNSIGSLSLRPQPLTFSPWGGFKSLFVPAPPGVEGTFSAP
PQRNRKSVHRVLAELDDESELENFPFVLMPEPRKLRVDKAFITPTGNRRGRPRKYP
VSAPMAPPAVGGGEPAAAPCCCLPQETVAVWVQCDGCDVWFHVACVGCSTQAAREADP
RCPGCRAGIQT"SEC ID N°: 4
RECUENTO DE BASES 795 a 774 c 791 g 661 t
ORIGEN
1 tcgtctcctg ccgaggagcc caagggggtcc cgggatccgc cgcacaggct ggcactgctt
61 gaagaggagg ctactcggag actgcgcgcg cgggattgga ggctccocaa atatttttgg
121 atctgaggat ccagctcaag tgaggtgcca taggacgtgt tcctgagttt gcattgcaog
181 gagacottoc tggaaattttt catttgcaag tcggcttaac caattttgca ttgagtccca
241 ggcctgcttc actotgaatt tgggctatcc aggtagtgtg ctcaaagttg aaaccgcata
301 cagcacaact caegtttgca tcagactggg aagcgaactt aagccagcgg tgcgtggccc
361 aggagtggga aaggaaatgg atgcctgaag tggagagggt ggtgcagagg gggcacogcc
421 catgtgcgcc tgcctccaac tgcctgcgat agggggcgcc agggggcggt atctctacac
481 cttccaoccc ccogcggggg ctggctgcac ctatcgcttg ggcacagggg ccgacctgtg
541 tgatgtggcc ctgcggcccc agcaggagcc tggcctcaco tctgggaccc accgocgaact
601 gcatgcogag ccccggggtg atgactggag ggtcagcctg gaagaccaca gcagccaagg
661 taotttggtc aataatgtcc gactcccaag aggtcacagg ctggaattga gtgatggaga
721 cctcctgacc tttggccctg aaggggccccc aggaaccagg ccctcggagt tctaactcat
781 gttccaacaa gtacgagtc agcctcagga ctttgcctgc attaccatcc caoggtotag
841 gggagaagcc cgggttgggg ctggtttccg gctatgctg cctccaccag gggctccaca
901 ggggctctct agcaccttct cccctgcgcc caaggccaca ctgatactaa actccatagg
961 cagcctcagc aagctccggc cccagccctt cactctctcc cctagtggg gtggaccaaa
1021 gagcctgcct gttcccgccc caactgggga agtggggaac accgcttctg ctccaccaca
1081 acgcaatcgg aggaatctg ttcacagagt gttgggggaa ctggatgatg agagtggcc
1141 tottgagaac ccgccaccgg tccctatgga gccagggaag aaactccgtg tagacaaagc
1201 cccactgact cccactggaa atcgactgg cctcctcgg aagtaaccag tgagcgtccc
1261 catggctccc cctgcagtgg ggggcgggga gccctgtgca gctccttgtt gctgctgccc
1321 ccagggaagag acagtggcct ggggttcagt tgatggctgt gacgtotgtt tccatgtgac
1381 ctgtgttggc tgcagcatcc aggetgocag ggaggocgac ttcagatgac cagggtgccc
1441 ggotggcatt cagaactaag gtccaccgcc aaggcaaat ccgacacacc tggccatgag
1501 tagacacago agogagcaaa taggtctgat aaataccccc cttcccttcc ctcccaggga
1561 gggaaatgact acagggaaga aggatggatt gatgtggact cattcagggc ctggagcaga
1621 cccgtgtggc caagacagaa gagatggttt cctgccaaag atattggccc ttcaggaaa
1681 ttgccagtga gctggaagt ccactatta caagccataa ggccatgttg ccatggacac
1741 cagaatatct gtagtcagag cactatcag ttgcaaaagc catgctgca accgatggaa
1801 aatgtaagag ggagttctta aggttcttga tggcatcacc caaggcattc tgggaaaaac
1861 tagggcctgg ccccaaaact tccctactct gggctagtc ctgctgocaa caaatcgtc
1921 ggcacctggc ttttcacagc tttgttttta tttcaagtc aaggacaagc cgttccattc
1981 actcctgggc atttactctt cttgtgggtc tgtgatattc cttgctttcc agggagaatg
2041 tgcttggcaa ggtctggaga actaatctag aatcttaggg gaagggggga gatggaaata
2101 caaacctgct tactggaaag gtgcaaatat atgggttgag ctggaggtag gaatacaggt
2161 aattaagggt tctggtttta gggaaaaacg atctattgcc atttaataaa ggttaactgg
2221 atttggttaa gttcacaaag ttagcagaag atttatttac aggottcacc tgaotgtca
2281 gggcaagaga aagcctggta aaccagctac agcagtttac cagtgtgatg gctgtgacac
2341 agotccactc caccgggtga caccagagag ggcaactggg ctggcctggt tcaigtgtga
2401 tcaaacogot taacccacac atggtacatg tgattttctt ttgtgagcct tacaccaagc
2461 caaactattg tcaaacogct atttotatag aaataaagcc ttatcttgac ctgttctatt

```

```

2521 aaaacctgoc acatccgcc tttcctacct agatttaatg agcccaagtt tttttacatg
2581 gaagaaatga ctctggggca aagaccccta atgaactagt ggcagagcoa ggaataaaac
2641 ttgagtaact aatgagtcac ttatgggcag agtatgcaa aacottaagt ggaaaccaa
2701 tagaccctgg tatcaagaaa gcacaaagta ttaatagaag tttctggttg gggatgatcta
2761 ggttcaacag aaataagatg atttctaagt ataaagccat tttagaattc cagagtaggg
2821 tgggaaagca aaaagccagc tctgaacagg taacagctac atggtgactg agtctatggg
2881 caaaagttot tgcatacacag gottttggga actagcctat cacagggccc tgtacaaata
2941 aacttggetg caatcccaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
3001 aaaaaaaaaa aaaaaaaaaa a SEC ID N°: 21

```

LOCUS	TCF 19	1080 bp	ARNm	lineal	PRI 05-NOV-2002
DEFINICIÓN	Factor de transcripción 19 de Homo sapiens (SC1) (TCF19) , ARNm.				
REFERENCIA	NM_007109				
VERSIÓN	NM_007109.1 GI:6005891'				
PAL. CLAVE	.				
FUENTE	Homo sapiens (humano)				
ORGANISMO	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCIA	1 (sitios)				
AUTORES	Ku,D.H., Chang,C.D., Koniecki,J., Cannizzaro,L.A., Boghosian-Sell,L., Alder,H. y Baserga,R.				
TÍTULO	A new growth-regulated complementary DNA with the sequence of a putative trans-activating factor				
PUBLICACIÓN	Cell Growth Differ. 2 (4) , 179-186 (1991)				
MEDLINE	<u>91329275</u>				
PUBMED	<u>1868030</u>				
REFERENCIA	2 (bases 1 a 1080)				
AUTORES	Krishnan,B.R., Jamry,I.. y Chaplin,D.D.				
TÍTULO	Feature mapping of the HLA class I region: localization of the POU5F1 and TCF19 genes •.				
PUBLICACIÓN	Genomics 30 (1) , 53-58 (1995)				
MEDLINE	<u>96129301</u>				
PUBMED	<u>8595903</u>				
REFERENCIA	3 (bases 1 a 1080)				
AUTORES	Kriehnan,R.				
TÍTULO	Molecular mapping by transposon-based nested deletion sequencing: the SC1 gene maps near the HLA-C locus				
PUBLICACIÓN	No publicado				
COMENTARIO	<u>REFSEQ PROVISIONAL</u> : Este registro no se ha sometido a una revisión				
CARACTERÍSTICAS	fina por el NCBI. La referencia de secuencia procede de <u>U25826.1</u>				
fuentes	Localización/Cualificadores				
	1..1080				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="6"				
	/map="6p21.3"				
	/clone="YAC B209D7, Cósido 52, pDEL5.5"				
	/haplotypes=>A3 B8 C- DR3 DQw2 DRw52 y A29 Bw65 C- DR7 DQw2 DP4"				
	/sexo="masculino"				
	/linea celular="linfoblastoide B"				
	/tipo tejido="sangre"				
	/lib_clone="biblioteca de CGM1 de la Washington University School of Medicine"				
	/estado_des="adult"				
gen	1..1080				

/gen="TCF19".
 /nota="sinónimo: SC1"
 /db_xref="locusID: 6941"
 /db_xref="MIM: 600912"
CDS 1..1080
 /gen="TCF19"
 /nota="factor de transcripción similar #"
 /codon_inicio=>l
 /producto="factor de transcripción 19 (SC1)"
 /id_proteína="NP_009040.1"
 /db_xref="GI: 6005892"
 /db_xref="LocusID: 6941"
 /db_xref="MIM: 600912"
 /traducción = "MLPCFOLLRIGGGGGDLYTFHPPAGVGCTYRLGHRADLCDVAL
 RPQQEPGLISGIAELHAEPRGDDWRVSLDEHSSQGLVNNVRLPRGHRLELSDGDL
 TFGPEGPPGTSPSEFYFMFQQVAVKPDFAITIPRSRGEARVAGGFRPMLFSQGA
 RPLSTFSPAPKATLIINSIGSLKLRPQPLTFSPSWGGPKSLFVPAPPGEVGTTPSAP
 PQNRNRKSVHRVLAELDDESEPPENPPVLMEPARKKLAVDKAPLTPGNARGRPKYP
 VSAEMAPPVAGAGSPVQLLVACPRKRQWPGFSVMVTSGSMWVPLAAASRLPGRPTS
 DAQGAGLAESLRSTAKAPSDTPAHE" SEC ID N°: 5

misc feature 88..264
 /gen="TCF19"
 /nota="FHA; Región: dominio asociado a Forkhead"
 /db_xref="CDD: smart00240"
misc feature 91..306
 /gen="TCF19"
 /nota="FHA; Región: dominio FHA. El dominio FHA
 (asociado a Forkhaad) es un motivo de unión a fosfopéptido"
 /db_xref="CDD: pfam00498"
conflicto 80
 /gen="TCF19"
 /citación=[1]
 /reemplazar="c"
variación 80
 /gen="TCF19"
 /nota="AVISO: mapa de localización ambiguo"
 /alelo="T"
 /alelo="C"
 /db_xref="dbSNP: 1065459"
conflicto 367..369
 /gen="TCF19"
 /citación=[1]
 /reemplazar="aac"
conflicto 577
 /gen="TCF19"
 /citación=[1]
 /reemplazar="a"
variación 631
 /gen="TCF19"
 /nota="AVISO: mapa de localización ambiguo"
 /alelo="G"
 /alelo="A"
 /db_xref="dbSNP: 2073721"
variación 657
 /gen="TCF19"
 /nota="AVISO: mapa de localización ambiguo"
 /alelo="C"
 /alelo="A"
 /db_xret="dbSNP: 2073722"

```

variación      722
                  /gen="TCF19"
                  /nota="AVISO: mapa de localización ambiguo"
                  /alelo="T"
                  /alelo="C"
                  /db_xref="dbSNP:2073724"
conflicto      8027..803
                  /gen="TCF19"
                  /citación=[1]
                  /reemplazar="cg"
conflicto      869..871
                  /gen="TCF19"
                  /citación=>[1]
                  /reemplazar=""
variación      1042
                  /gen="TCF19"
                  /nota="AVISO: mapa de localización ambiguo"
                  /alelo="G"
                  /alelo="A"
                  /db_xref="dbSNP:1065460"
conflicto      1045..1046
                  /gen="TCF19"
                  /citación=[1]
                  /reemplazar="cg"
RECuento de base 200 a 361 c 315 g 204 t
ORIGEN
1 atgctgccct gcttccaact gctgcgcata gggggcggca ggggcgggtga tctctacacc
61 ttccacccccc ccgcgggggt tggctgcacc tatcgcttgg gccacagggo cgacctgtgt
121 gatgtggccc tgcggcccca gcaggagcct ggootcatct ctgggatcca cgcogaactg
181 catgccgagc ccgggggtga tgactggagg gtcagcctgg aagaccacag cagccaaggt
241 acttttggtca ataattgtccg actcccaaga ggtcacaggc tggaaattgag tgatggagac
301 ctectgacct ttggccctga agggcccccga ggaaccagcc cctcggagtt ctacttcatg
361 ttccaacaag tacgagtcaa gcctcaggac tttgctgcca ttaccatccc acgggtctagg
421 ggagaagccc gggttggggc tggtttcogg cctatgctgc cctcccaggg ggctccacag
481 cggcctctca gcacettctc ccctgcccc aaggccacac tgatootaaa ctocataggo
541 agcctcagca agctcoggoc ccagcccctc acottctccc ctagtggggg tggaccaaaag
601 agcctgcctg ttcccgccoc acctggggaa gtggggacca cgccttctgc tccaccccaa
661 cgcantcgga ggaatatgt tcaccgagtg ttggcggaac tggatgatga gagtgagcct
721 cctgagaacc cgcacccggt ccttatggag ccaggaaga aactccgtgt agacaaagcc
781 ccaactgactc ccaactggaaa tgcacgtggc cgtcctcgga agtaccaggt gagcgctcac
841 atggctcccc ctgcagttgg ggcggggagc cctgtgcagc tccttgttgc tgcctgccc
901 aggaagagac agtggcctgg gttaagtgtg atggotgtga cgtctgggtc catgtggcct
961 gtgttggtg cagcatccag gctgcccagg aggcgcactt ccgatgccc ggggtgccggg
1021 ctggcattea gcctaaggtc caccgccaag gcaocatcgg acacacctgc ccatgagtag
SEC ID N°: 22

```




PubMed Nucleótido Proteína Genoma Estructura PMC Taxonomía OMIM Boo

Buscar para Ir

Limites Vista previa/índice Historia Portapapeles Detalles

Presentar 20

1: AB029519. Homo sapiens SC1(...[gi:5420470]) [Links](#)

LOCUS AB029519 . 4651 pb ADN lineal PRI 29-APR-2000
 DEFINICIÓN Gen de SC1-7 (TCF19) de Homo sapiens, cds completo.
 REFERENCIA AB029519
 VERSIÓN AB029519.1 GI:5420470
 PAL. CLAVE
 FUENTE Homo sapiens (humano)
 ORGANISMO Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; -Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCIA 1 (sitios)
 AUTORES Teraoka, Y., -Naruse, T.K.; Oka, A.; Matsuzawa, Y., Shiina, T.,
 Iizuka, M., Iwashita, K., Ozawa, A. e Inoko, H.
 TÍTULO Genetic polymorphisms in the cell growth regulated gene, SC1
 telomeric of the HLA-C gen and lack of association of psoriasis
 vulgaris
 PUBLICACIÓN Tissue Antigens 55 (3), 206-211 (2000)
 MEDLINE 20236853
 REFERENCIA 2 (bases 1 a 4651)
 AUTORES Teraoka, Y. e Inoko, H.
 TÍTULO Remisión directa
 PUBLICACIÓN Remitido (02-JUL-1999), Yoshika Teraoka, Tokai University School of
 Medicine, Molecular Life Science; Bohseidai, Isehara, Kanagawa
 259-1193/ Japón (E-mail.: ybshika@is.icc.u-tokai.ac.jp,
 Tel: 81-463-93-1121 (ex. 2.653), Fax: 81-463-94-8884)
 CARACTERÍSTICAS Localización/Cualificadores
 fuente 1..4651
 /organismo="Homo sapiens"
 /db_xref="taxon:9606"
gen unión(378..615,2320..2878,3346..3586)
 /gen="SC1-7 (TCF19) "
CDS Unión(378..615,2320..2878,3346..3586)
 /gen="SC1-7 (TCF19) "
 /codon inicio=1
 /id_proteína="BAA82327.1"
 /db_xref="GI:5420471"
 /traducción = "MLPCQLLRIGGGRGGLYTFHPPAGAGCTYRLGHRADLCDVAL
 RPQQEPGLISGIHAELEFRGDDWRVSLSDHSSQGLVNNVRLPRGHRLELSDGDL
 TFGPEGPPGTSPPSEFYFMFQQVRVKPDFAAITIPSRGEARVAGAGFRMLESQCAPQ
 RPLSTFSPAPKATLIINSIGSLSKLRPQLTFSPSWGGPKSLVPEAPPGEVGTTPSAP
 PQNRNRKSVHRVLAELDDESEFPENPFVLMPEPRKKLRVDKAPLTPGNRRGRPKYP
 VSAPMAPPAVGGGEPCAPCCCLPQEBTVAVVQCDGCDVWFHVACVGCISQAAREADF
 RCPGCRAGTQT" SEC ID N°: 6
 RECUENTO DE BASES 1160 a 1163 c 1244 g 1084 t
 ORIGEN
 1 ccgcaccccc gacacacaca ggagctgcct aaagtatcct tgcottgcag attggaggct
 61 ccccaaatat tttgtgatct gaggatccag ctcaagtgcg gtgccatagg acgtgttcct
 121 gagtttgcct tgcacggaga ccttcctgga atttttcatt tgcaagtcgg ottaaccaat

181 ttgcatga gtcctaggct gettgaetc fgaatttggg ctattcaggt agtgtgetca
 241 aagttgaaac cgcatacagc acaactcaag tttgcatcag actgggaagc gaacttaagc
 301 cagcgggtgcg tggcccagga gtgggaaagg aaatggatgc ctgaagtggg agaggttggtg
 361 cagaggggggc accgcccagc ctgcccgtct tccaactgct gcgcataggg ggcggcaggg
 421 gcggtgatct ctacaccttc ccccccccg cgggggtggt ctgcacctat cgcttggggc
 481 acagggccga cctgtgtgat gtggccctgc ggcgccagca ggagcctggc ctcatctctg
 541 ggatccacgc cgaactgcat gcagagcccc ggggtgatga ctggaggggtc agcctggagc
 601 accacagcag ccaaggtgag cattaagcag ggcaggtttg cccctgggtg gttgaagcgc
 661 cagggtggaa tgagttaagg ctccacaaga cccctgtgoc tgcctcccat actccatca
 721 gattggatgg atagtctgtg tccagacctt catcttccca ccagaagtgt gcacagtacg
 781 aagctctctg ccagactgac cctttttggt cccgtttagc tcatacagga cctgggatat
 841 catcagaaag atatcacagt ggggatgttc tgaggccact agaggccaaag tttagacttg
 901 attcagtttc cagctttgct gaggcactct gttcctgggt tagggcagtt ctatgttgaa
 961 taatgttttt aataatctgg gcatgtottt ctccgtgact tgaggcagtt agcctcagaa
 1021 agcctagatt oacatttgag ttttgccact gectctgggt aaagtcaagt gtagagtggt
 1081 tatggttatt agactatagt agcacaact catctagtgc ttactgttat gagccaggcc
 1141 ctatttttaag tgtattgaat gtggtgggtg ctaatatatt cctcatttac agtaaaaggaa
 1201 aatgaggcac aaagaggtta aggaacttgt ccagggtgtg goatgggtgtg ttacacotat
 1261 aatccagcac tttgggaggg taaggcaggg tggatcactt gagctcagga gttcagaccc
 1321 agcctgggca acatggtgaa aaactgtctc taccaaaaaa ttaattaatt ttttaaaaaa
 1381 agcctgggag cgggtgtgca cgcctgtaat pccagcactt tgggagggcg agatgggcag
 1441 atcaagaggt caggagttcg agaaccatct gaccaacatg ttgaancccc atctgtgctg
 1501 aaaaaaaat acaaaaatta gccaggtgtg gtggcgtgca cctgtaacco cagctactca
 1561 ggaggtgaa gcagcagaat cacttgaaac cgggaggggg aggttgcagt gagotgagat
 1621 cgaacactg caotocagct tgggagagag agcagagctc oatotoaaac aaacaaacaa
 1681 accaaaagct tgcccagggg acataactg gtaagtggta gagctaggat ctgaacgagc
 1741 tggatgtggg ggagagtga catgtttgaa aactggacct tagggcgggg caoggtggt
 1801 cagcctgtg atccagcac tttgggaggg tgaggogggc agatcaggag gtcaggagta
 1861 tgagaccagc ctggocaaac tggtaaaadcc ctgtctctgc taaaaataaa aaaattagcc
 1921 agcgtgtgtg gcacatgcct gtaateccag ctactcagga ggctgaggca ggagaattgc
 1981 ttgaacotgg gaggcgggtc aagcttgggc aatagagcaa aactocactc caaaaaaaa
 2041 aaaaagaaag aaaaaaaag aagaaagaa gaaaattgga ccttaggaca gtgagggcag
 2101 ggatootttg taggaagca caagaaacac agactgttc cttagtgaca aggagtgtac
 2161 tgctgtgtac ctgtcacctg ctgaggggtc taggatgtga gggagaatct gactacagtt
 2221 tcataattct cccagaaaat catacagatt tctccactcc tgaactctgtt catttctgtt
 2281 tttgtcctcc atatttgcct ggtgcccac catcaacagg taotttgggt aataatgtcc
 2341 gactocaaag aggtcacagg ctggaattga gtgatggaga cctcctgacc tttggccctg
 2401 aagggccccc aggaadacag cctcggagt totacttcat gttocaaaca gttcagagtc
 2461 agcctcagga ctttgcctgc attacatcc caggtctag cggagaagcc cgggttgggg
 2521 ctggtttccg gcctatgctg cctccacagg gggctccaca gggcctctc agcaccttct
 2581 cccctgcccc caaggccaca ctgatccata actccatagg cagcctcagc aagctccggc
 2641 cccagccctc caccctctcc cotagttggg gtggaccaaa gagcctgcct gttccggccc
 2701 caactgggga agtggggacc aogccttctg ctocaccoca acgcaatcgg aggaatctg
 2761 ttcacogagt gttggcgga atggatgat agagttagcc tcttgagaac ccgcaaccgg
 2821 tocttatgga gccaggaag aaactcctg tagacaaag cccactgact cccactgggt
 2881 aagtggagtc ctacttggc cctctcagt ttttactgt tttogattcc ttgtatccct
 2941 aggtgtgag gaggtccopg tgcotggggg gatgggcacg ggaggtggaa tagatggaa
 3001 ggcaagacct gggtagctc tgataggaag agaaaaatat gtgcaggaga acatgagag
 3061 tggggtgggg cagtgttat aaaaacacc gagttagcat gtccgtcttt tcacattcat
 3121 atggctttta cccattctt ctagtcccta aggatgggga actttcaggt tcaactaga
 3181 ggttttttag cccaccotat gtgtttttaa ggacagagtc caggctcacc ttagtctca
 3241 gaccactgtg cctctgtggc ctacccctat gaccagccat aggggtggca ggtotaggcc
 3301 tctcctaca ggtttccggg gacccctgtg tctgtgtcac ttcagaaatc gacgtggcgg
 3361 tctcgggaag tacocagtga gcgtcccat ggctccccc gcaagtgggg gcggggagcc
 3421 ctgtgcagct ccttgttgc cctgccccg ggaagagaca gtggcctggg ttcaggtgga
 3481 tggctgtgac gtctgttcc atgtggctg tgttggctgc agcatccagg ctgccggga
 3541 ggccgacttc cgtgcccag ggtgcccggc tggcattcag acctaaagtc caccgccaag
 3601 gcaccatcgg acacacctgc ccatgagtag acacagcagc gagcaaatag gtotgataaa
 3661 taccacctt cccctccctc ocaagaggg aatgaetaca gggaaagagg atggattgat
 3721 gtggactcat toagggcctg gacagaccc tgggtggcaa gacagaagag atggtttcct


```

3781 gccaaagata ttgccacctc caggaaattg ccagtgaget ggaagttccc actattacaa
3841 gccataaggg catgttgcca tggacaccag aatatctgta gtcagagcac ctatcagttg
3901 caaaagccat gcotgcaacc gatggaaaat gtaagaggga gttcttaagg ttcttggtgg
3961 catcaccocaa ggcattctgg gaaaacctag ggcctggccc caaaacttcc ctactctgtg
4021 gctagtcctg ctgccaacaa aatcgtagcg acctggcttt tcacagcttt gcttttattt
4081 ocaagtcaag gacaagccgc ttcattcact cctgggcatt tactcttctt gtgggtotgt
4141 gatattcctt gctttccagg gagaatgtgc ttggcaaggc ctggagaact aattcagaat
4201 cttaggggaa ggggagagat ggaaatacaa acctgcttac tggaaaggty caaatatatg
4261 gggtgagctg gaggtaggaa tacaggtaat taaggtttct agtttaaggg aaaacagatc
4321 tattgccatt taaataaggt aactgggatt tggtaagtt cacaagata gcagaagatt
4381 tatttaacagg cttcacctgt actgtcaggc caagagaaag cctggtaaac cagctacagc
4441 agtttaccag tgtgatggct gtgacacagc tccactccac ggggtggcac agcagagggc
4501 aactgggctg gcctgggtca gtgtgaatca aacggcttaa cccacacatg gtacatgtga
4561 ttttcttttg tgagccttac accaagcoaa actattgtca aagcatcatt tctatagaaa
4621 taaagcotta tcttgacctg ttctattaaa a SEC ID N°: 23

```




[PubMed](#) [Nucleótido](#) [Proteína](#) [Genoma](#) [Estructura](#) [PMC](#) [Taxonomía](#) [OMIM](#) [Boo](#)

Buscar para

[Limites](#) [Vista previa/índice](#) [Historia](#) [Portapapeles](#) [Detalles](#)

1: U25826. Human transcripti...[gi:833832] [Links](#)

LOCUS	HSU25826	5522 pb	ADN	lineal	PRI 27-MAY-1995
DEFINICIÓN	Gen de factor de transcripción (SC1) humano, cds completo.				
REFERENCIA.	U25826				
VERSIÓN	U025826.1 GI:833832				
PAL. CLAVE					
FUENTE	Homo sapiens (humano)				
ORGANISMO	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; 'Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCIA	1 (bases 1 a 5522)				
AUTORES	Krishnan,R.				
TÍTULO	Molecular mapping by ,transposon-based nested deletion sequencing: the SC1 gen maps near the HLA-C locus				
PUBLICACIÓN	No publicado				
REFERENCIA	2 (sitios)				
AUTORES	Ku,D.H., Cnang,C.D., Koniecki,J., Cannizzaro,L.A., Boghosian-Sell,L., Alder, H. y Baserga; R.				
TÍTULO	A new growth-regulated complementary DNA with the sequence of a putative trans-activating factor				
PUBLICACIÓN	Cell Growth Differ. 2 (4), 179-186 (1991)				
MEDLINE	91329275				
PUBMED	1868030				
REFERENCIA	3 (bases 1 a 5522)				
AUTORES	Krishnan,R.				
TÍTULO	Remisión directa				
PUBLICACIÓN	Remitido (27-ABR-1995) Rajendra Krishnan, Dept. of Internal Medicine, Washington University School of Medicine, Div. Allergy & Immun., 660 South Euclid Avenue, St. Louis, MO 63110, USA				
CARACTERÍSTICAS	Localización/Cualificadores				
fuelle	1..5522 /organismo="Homo sapiens" /db_xref="taxon:9606" /cromosoma="6" /mapa="6p23.1" /clon="YAC B209D7, Cósmido 52, pDEL5.5" /haplotipo="A3 B8 C- DR3 DQW2 DRw52 y A29 Bw65 C- DR7 DQw2 DP4" /sexo="masculino" /línea_celular="linfoblastoide B" /tipo_tejido="sangre" /lib_clon="Biblioteca de CGM1 de la Washington University School of Medicine" /est_test="adulto" unión(933..1170,2905..3463,3929..4211) /gen="SC1" unión(933..1170,2905..3463,3929..4211) /gen="SC1"				
<u>gen</u>	unión(933..1170,2905..3463,3929..4211)				
<u>CDS</u>	unión(933..1170,2905..3463,3929..4211)				

```

/codon_inicio=1
/producto="factor de transcripción SC1"
/id_proteina="AAB60363.1"
/db_xref="GI: 833833"
/ traducción ="MLPCFQLLRIGGGGGDLYTFHFEAGVGCTYRLGHRADLCDAVAL
RPOQEPGLISGIHAELHAEPRGDDWRVSLDESSQGTLVNNVRLPRGHRLELSDGDL
TFGPEGPPGTSPSEFYFMFQQVRVKPQDFAAITIPRSRGEARVAGFRPMLPSQGAPO
RPLSTFSPAPKATLILNSIGSLSKLRPOPLTFSPSWGGPKSLPVPAPGGEVGTTPSAP
PQRNRKSVHRVLAELDDESEPPENPPPVLMEPRKKLRVDKAPLTPTGNARGRPKYP
VSAPMAPPAVGAGSPVQLLVAAACPRKROWPGFSVMAVTSGSMWFLAAASRLPGRPTS
DAQGAGLAFSLRSTAKAPSDTEAHE" SEC ID N°: 7

exón          <933.1170
               /gen="SC1"
conflicto     1012.
               /gen="SC1"
               /citación=[2]
               /reemplazar="c"
intron        1171..2904
               /gen="SC1"
Región repetición 1937..2238
               /familia_rpt="Alu"
exón          2906..3463
               /gen="SC1"
conflicto     3033..3035
               /gen="SC1"
               /citación=[2]
               /reemplazar="aac"
conflicto     3243
               /gen="SC1"
               /citación=[2]
               /reemplazar="a"
intron        3464..3928
               /gen="SC1"
exón          3929..>4211
               /gen="SC1"
conflicto     3933..3934
               /gen="SC1"
               /citación=[2]
               /reemplazar="cg"
conflicto     4000..4002
               /gen="SC1"
               /citación=[2]
               /reemplazar=""
conflicto     4176..4177
               /gen="SC1"
               /citación=[2]
               /reemplazar="cg"

RECuento DE BASES 1376a      1368 c      1497 g      1281 t
ORIGEN

1  gaattccgaa ggaggggtag ggcgtgccg cgcgcagagg ccgcgccct cctggcccg
61  gctttcttgc tgcacaacag tagcagcaac gtcggctcct gccgaggago ccaaggggtc
121 cccggatccg ccgcacaggc tggcactgct tgaagaggag gctactcgga gactgcgccg
181 ccggggtaga tccgaaacgg ggcctggggc gagtgggaaa aggcggggtg tgccttgcct
241 gatcgogggg agctccttcc tgtttttatc ccaoctagag aagccgggaa gtagggggtt
301 aggtccaatt tgttggaata cttaaggact cgtttgcact ttcttttggg ggatgacagt
361 ggattcattg ccctoggagg ttcaaccagt tatgagttag ggattggcca gaagatccgg
421 gccgcagga gacaggagtgc tctattagga taagcaagtt tgacagggaag aagctgctct
481 tctccgaatt acacagaggt gatgtgttog tattgcaggt agcgtgtgt ataacaggac
541 ctctttcccc gcgcgccgcc accccgacac acacaggagc tgcctaaagt atccttgcct
601 tgaagattgg aggcctccca aatatttggg gatctgagga tccagctcaa gtgaggtgcc

```

661 ataggacgtg ttcctgagtt tgcattgoc ggagacotto ctggaatttt tcatittgcaa
 721 gtccggcttaa ccaatttggc attgagtcct aggotgottg cactctgaat ttgggctatt
 781 caggtagtgt gtocaaagtt gagacogcat acagaacaac toaagtttgc atcagactgg
 841 gaagcgaact taagccagcg gtgcgtggcc caggagtggg aaaggaatg gatgcctgaa
 901 gtggaagagg tgggtgcagag ggggcaccgc coactgtgcc ctgcttccaa ctgctgcgca
 961 tagggggcgg cagggggcgg gatctctaca ccttccaccc ccccgccggg gttggctgca
 1021 cctatcgctt gggccacagg gccgacctgt gtgatgtggc cctgogggcc cagcaggagc
 1081 ctggccctcat ctctgggac caccgccaac tgcctgcoga gccccggggg gatgactgga
 1141 gggtcagcct ggaagacac agcagccaa gtgagcatta agcagggcag ctttgcocct
 1201 ggggtggtga agcgcagggc tggaaatgag aaggtctcca caagaocctg ctgtctgcct
 1261 cccatactoc catcagattg gatggatagt ogtggtccag accttoatct tccaccaga
 1321 agtctgcaca gtcagaagct ctctgccaga ctgacctttt ttggtccogt ttgactcata
 1381 caggacctgg gatctcatca gaagatata acagtgggga tgttctgagg ccactagagg
 1441 ccaagtttag acttgattca gtttccagct ttgctgaggg actctgttcc tgggttaggg
 1501 cagttctatg ttgaataatg tttttaataa totgggcatg tcttctccg tgacttaggg
 1561 cagttagcct cagaaagcct agattccat ttgagtttg ccactgccto ttggtaaaagt
 1621 cagctgtagg agtgttatgg ttatttagact atagttagca acattcatct agtgcctact
 1681 gttatgagcc aggcocctatt ttaagtgtat tgaatgtagg tggtaactat attatcctca
 1741 ttacagaaa aggaatga ggcacaaaga ggttaaggaa cttgtccagg gctgggcatg
 1801 gtggtttaca cctataatcc agcactttgg gaggctaagg caggggtggat cacttagct
 1861 caggagttcg agaccagcct gggcaacatg gtgaaaacct gtctctacca aaaaattaat
 1921 tagtttttta aaaaaagcct gggcggcgtg ggtcagcct gtaatccag cactttggga
 1981 ggcagagatg ggcagatcac gaggtoagga gttcagagcc atcctgacca acatgttgaa
 2041 accccatctg tgotgaaaaa aaaaaaccca aattaaocag gtgtgtggc gtgcacctgt
 2101 aacccagct actcaggag ctgaagcagc agaatcaott gaacocggga ggcggaggtt
 2161 ggagttagct gagatcgac cactgcacto cagcttgggc gacagagcga gactccatct
 2221 caaacacaca aacaaaccaa aagcttgccc agggtcacat aactggtaag ttgttagct
 2281 agtactgaa cggctggag ctgggggaga gtgagcatgt ttgaaaactg gacottaggg
 2341 cggggcacgg tccgtcacgc ctgtaatccc agcactttgg gaggctgagg cgggcagatc
 2401 agggagtcag gattatgaga ccagocctgg caacatggta aaacocctgt tctcgtaaaa
 2461 ataaaaaat tagccagagc tgggtggaca tgcctgtaat ccagatcact caggagctg
 2521 aggcaggaga attgcttgaa cctgggaggg ggagtgcagt gagctgagat tgcactactg
 2581 cactocagct tgggcaatag agcaaatctc catctcaaaa aaaaaaaaaa gaaagaaaaa
 2641 aaagaagaa agaaagaatt ttggccotta ggacagttag ggcaggggtc ctttgtggga
 2701 aagcacaaga aacacagatt tgttctagc tgacaaggag tgtactgcct ggtacctgtc
 2761 acctgctgag ggccttagga tgtgaggag aatctgaoca cagtttcata ttcttccca
 2821 gaactatc agatttctcc actcctgact ctggtcaatt ctgttttct cctccatatt
 2881 tgcctggtgc cccaccatca acaggtactt tggtoataa tgtccgactc ccaagaggtc
 2941 acaggctgga attgagtgat ggaagacotc tgaoccttgg cccctgaagg ccccccaggaa
 3001 ccagccctc ggagtctcat ttcatgttcc aacaagtacg agtcaagcct caggactttg
 3061 ctgocattac catcccacgg tctaggggag aagocccggg tggggctggg ttccggccta
 3121 tgotgocctc ccagggggct cccagcgggc ctctcagcac cttctccctc gcccacagg
 3181 ccacactgat cctaaactcc ataggcagcc tcagcaagct cgggcccag cccctccct
 3241 tctccocctag ttgggttgga ccaagagac tgcctgttcc cggcccaoct ggggaagtg
 3301 ggaocacgoc ttctgttcaa ccccaagga atcggaggaa atctgttcao cgagtgttg
 3361 cggaaactgga tgatgagagt gacocctctg agaacccgcc accggtcctt atggagccca
 3421 ggaagaaact ccgtgtagac aaagccacac tgactccac tgggtaagt ggtctctac
 3481 ttggccctct cagtgttcta ctgcttttgg attccttgta tccctaggct gtgaggaggt
 3541 cccocctgct ggggggatgg gcacgggag tggaaatgat ggaatggcaa gacctgggtt
 3601 agctctgata gggaaagaaa aatatgtgca ggagaacatg agaggtgggg tggggcagtg
 3661 ctataaaaca accggagtg gcattgtcctg ctttttcat tcatatggct ttaacccac
 3721 tttctagtgc ctaaggatgg ggaacttca ggctcacact agaggttttt agggccaccc
 3781 catgtgtttt taaggaoaga gtcaggctc accttagtto tcagaccact gtgcctctgt
 3841 ggocctcaccc tatgaccago catagggtgg caaggcttag gccttctcag atttccgggt
 3901 accotttgtt ctctctcact tcttoagaa atgcaogtgg ccgtcctcgg aagtaaccag
 3961 tgagocctcc catggctccc cctgcagttg ggggggggag ccoctgtgca ctocctgttg
 4021 ctgocctgcc caggagaga cagtggcctg ggttcagtg gatggctgt agctctggtt
 4081 ccatgtggcc tgtgttggt gcagcatcca ggctgccagg gaggccagct tccgatgccc
 4141 aggggtgcgg gctggcattc agcctaaggt ccacogccaa ggcacacatg gacacacctg
 4201 cccatgagta gacacagcag cgagcaataa ggtotgataa ataocccoct tcccttccct

```

4261 ccccaagagg gaatgactac aggggaagaag gatggattga tgtggactca ttcagggcct
4321 ggagcagacc ctggtggcca agacagaaga gatggtttcc ttccaaagat attgccacct
4381 ccaggaaatt gccagtgaagc tggaaagtcc caotattaca agccataagg ccatgtcgcc
4441 atggacacca gaatatctgt agtcagagca cctatcagtt gcaaaagcca tgcctgcaac
4501 cgatggaaaa tgtaagaggg agttcttaag gttcttggtg gaatoaccca aggtattctg
4561 ggaaaacctc gggcctggcc ccaaaacttc cctactctgt ggctagtcct gctgccaaca
4621 aaatcgtagc gacctggctt ttcacagctt tgcttttatt tccaagtcaa ggacaagocg
4681 cttcattcac tectgggcat ttactcttct tgtgggtctg tgatattcct tgctttcag
4741 ggagaatgtg ctgggcaagg tctggagAAC taattcagaa tcttagggga aggggagaga
4801 tggaaataca aacctgctta ctggaaaggT gcaaaatag ggttgagctg gaggtaggaa
4861 tacaggtaat taaggtttct agtttaaggg aaaacagatc tttttgcat ttaaatatgg
4921 taactgggat ttggttaagt tcaccagat agcagaagat ttatttacag gcttcacctg
4981 tactgtcagg gacaagagaa aagcctggta aaccaagcta cagcagttta ccagtgtgat
5041 ggotctcaca cagctccacc ccccggttg acacagcaga gggcacctgg gctggcctgg
5101 ttcagtgtga atcaaacgc ttaaccaca catggtacat gtgattttct tttgtgagcc
5161 ttacaccaag ccaaaactatt gtcaaacgat catttctata gaaataaagc cttatcttga
5221 cctgttctat taaaacctgc cacatccgoc ctttctacc tagatttaat gagcccaagt
5281 ttttttacat ggaagaaatg actctggggc aaagaccctc aatgaactag tggcagagcc
5341 aggaataaaa cttgagtaac taatgagtca cttatgggca gagtatgcaa aaaccttaag
5401 tggaaaccaa atagaccctg gtatcaagaa agcaaaaagt attaatagaa gtttctggtt
5461 ggggtgatct aggttcaaca gaataaagat gatttctaag tataaagctc aaaattgaat
5521 tc SEC ID N°: 24

```

[illegible]

E1: NM_010097. Mus musculus SPAR...[gi:6753707]

Links[†]

174

IAGVEATADIENHPNDKAEKPSALNSEEETHEQSTEQDKTYSFEVDLKDEEDGGDLIS
 VDPTEGTLTLDLQEGTSEPQQKSLPENGDFPATVSTSYVDPNQRANITKCKESQEQPV
 SDSHQPNNESSKQTQDLKAEESQTQDPDIPNEEEEEEEEEEEEEEEEEEPEDIGAPSDNQ
 EEGKEPLEEQPTSKWEGNREQSDDTLESSQPTQISKTEKHQSEQGNQGOESDSEAE
 EDKASGSKEHIPHTEQDDQEGKAGLEAIGNQKDTDEKAVSTEPTDAAVVPRSHGGAGD
 NGGGDDSKHGAGDDYFIPSOEFLEAERMHSLSYLLKYGGGEETTTGESENQREAADNQ
 EAKKAESSPNAEPPSDEGNSREHSAGSCTNFQCKRGHICKTDEQGGKPHCVQDPPETCPP
 AKILDQACGTDNQTYASSCHLEATKCRLEGTKKSHQLQLDYFGACKSI PACTDFEVAQ
 FPLMRDWLKNILMQLYEPNPKHGGYLNKQKRSKVKKIYLDKRLLAGDHPIELLLRD
 FKINYHMYVYPVHWQFNELDQHPADRILTHSELAPLRASLVPMEHCITRFFEECDPNK
 DKHITLKEWGHCFGIKEEDIDENLLF" SEC ID N°: 8

misc feature 1573..1644
 /gen="Sparc11"
 /nota="FOLN; Región: Similar a dominio N-terminal de folistatina"
 /db_xref="CDD:smart00274"
misc feature 1645..1806
 /gen="Sparc11"
 /nota="kazal; Región: dominio inhibidor de serin proteasa de tipo kazal. Normalmente indicativo de inhibidores de serina proteasa.
 Sin embargo, los dominios de tipo kazal también se observan el aparte extracelular de las agrinas, que no se sabe que sean inhibidores de proteasas. Los dominios kazal a menudo aparecen en disposiciones en tándem. Plegamiento alfa+beta pequeño que contiene tres disulfuros. El alineamiento también incluye un solo dominio de transportadores en la familia de OATP/PGT"
 /db_xref="CDD:pfam00050"
misc feature 1654..1806
 /gen="Sparc11"
 /nota="KAZAL; Región: inhibidores de serin proteasa de tipo kazal"
 /db_xref="CDD:smart00280"

RECuento DE BASES 862 a 642 c 661 g 548 t
 ORIGEN

```

1  cagcacggag ggagcgagat ccaggaatct gcaacagaaa ccatgacagc ctgaaacacc
61  ctgtggtgcc aacctccaaa ttctcatctg tcacttcaga cootgactgg ctgacagagc
121 agcagaattt caactccaat aaactgtgaat gtgcttctag gcaagcagc caagtgcagc
181 agggaggggg gtggaagagc tagctoctct tgggcatttg tcaaaccttt caactcctggc
241 tgtgtgcaag gagggggactc aacttcggct tcaagctacc aaggctctgg atccagccac
301 ctctcogcag atotagcag catgaaggct gtgcttctcc tctgtgtgag cttgggaacc
361 gctgtggcaa tcccgacaag tacaagggtt ctctctgacc actccaacc aactactgca
421 acactgtgta caccggaaga cgtacagtc cccattgccc ggggtgaagc tacagcagac
481 atagaaaacc atcccaatga caaggctgaa aaacottcag caottaattc agaagaggaa
541 actcatgaac agtcaacaga gcaggacaaa acotacagct tcgagggtgga cctgaaggat
601 gaggaggatg gagatgggga ttttaagtga gatccaaagg aaggaaact aacaotggat
661 ctacaagaag gtacaagtga gcctcaacag aaaagtctcc cggagaacgg cgatttccc
721 ggcaccgtgt ccacttctta tgtggatcct aaccaacggc caaacatcac aaagggaaag
781 gagagtcagg agcaacotgt aagtgaacta caccagcaac cgaatgaaag cagcaagcaa
841 acccaagact taaaggotga agaaagccag acacaagatc cagacattcc caatgaagaa
901 gaggagaag aagaggaaga agaagaggaa gaagaggaa agccggaaga cattggtgcc
961 cccagtata accaagagga gggaaaagaa cctctggagg agcagcctac cagcaagtgg
1021 gaagyaaca gagagcaatc tgatgacacc ttagaagagt ccagtcagcc cactcagata
1081 agcaagacag agaagcatca atctgagcaa ggaaccaa ggcaggagag tgactctgag
1141 gcagaaggag aggaacaggc ttcaggcagc aaggaaacaa ttccacatac agagcagcag
1201 gaccaagaag ggaagctgg ccttgaagct attggcaacc agaaggacac tgatgagaag
1261 gccgtttcca cagaacotac cgaatgctgc gtggtgccta ggagtcacgg aggagctggt
1321 gataacgggg gcggggatga ctctaagcat ggtgcaggcg atgactactt catcccagc
1381 caggaattcc tagaggccga aaggatgcat tccctctcct attacctcaa atatggcgcc
1441 ggcgaggaga caacgactgg cgagagttag aaccagaggg aggctgcaga caaccaagag
1501 gccaaagaag ctgagagctc accaaatgot gaacottcag atgagggcaa ctcaaggag
1561 cacagtgtgt gttcttgcaac gaacttccaa tgtaaaaggg gacacatttg caaaacogat

```

```

1621 ccacaaggga aacotcactg tgtttgcaa gatccagaga cttgtccccc tgcaaaaatc
1681 otagatcagg ottgtggcac tgacaaccaa acctacgcca gctcctgtca cctgtttgct
1741 accaagtga ggctggaggg gaccaaaaag ggacaccaac tgcagctgga ttacttcgga
1801 gottgcaaat ctattcctgc ttgtacggac tttgaagtgg ctcagtttcc cctgoggatg
1861 agagactggc tcaaaaacat cctcatgcag ctttatgaac caaatcccaa acatggcggc
1921 tatctcaatg aaaagcaaaag aagcaaagtc aaaaaaattt acctggatga gaagagactc
1981 ttggtgggag accatcccat tgaacttctc ttgaggggact ttaagaaaaa otaccacatg
2041 tatgtgtatc ctgtgcaactg gcagtttaat gaactggacc agcatcctgc agacaggatc
2101 ttgacacact ctgaaottgc tctctcgoga gcttccctgg tgcccatgga acaactgcata
2161 actcgtctct ttgaggagtg tgaccocaaac aaggataagc acatcacctt gaagggaatg
2221 ggccactgot ttggaattaa agaggaggat atagatgaaa acctcctctt ttgaattag
2281 atttgagaga atcggaactt tccatocacc tcacctgctt taaccgcttc agaaatacga
2341 gcagccatga cactatacat tcatatgtag caaaacattt gtttggcatg tgagagaaga
2401 caatggtagt aattacttct tggatgata tatatgagcc aggoacttaa tattaactta
2461 ggaaatgaaa ctttaaaatt aagtagagtc aatgtctata aaagactgtc ctgtctgggg
2521 acagttagcc accatggcaa tgtactctg tgcactctgc tttataattg ataattataa
2581 actattaaaa aaacaatgtt catattgtcc ataatacctt atgcatgctg aggaagtga
2641 ataactgtct tttgagataa atatgcctcc ttttcagtgt cttgatgtcc taataaaaaa
2701 tctataaaac ccc SEC ID N°: 25

```





[PubMed](#) [Nucleótido](#) [Proteína](#) [Genoma](#) [Estructura](#) [PMC](#) [Taxonomía](#) [OMIM](#) [Boo](#)

Buscar para

[Limites](#) [Vista previa/índice](#) [Historia](#) [Portapapeles](#) [Detalles](#)

Fig 1: U64827. Mus musculus extr...[gi:1546922] [Links](#)

LOCUS	MMU64627	2734 pb	ARNm	lineal	ROD 30-NOV-2000
DEFINICIÓN	ARNm de la proteína asociada a la matriz extracelular (SC1) de Mus musculus, Cds completo.				
U	U64827				
VERSIÓN	U64827.1	GI:1546922			
PAL. CLAVE					
FUENTE	Mus musculus (ratón doméstico)				
ORGANISMO	<u>Mus musculus</u> Eukaryota; Metazoa," Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 a 2734)				
REFERENCIA	McKinnon, P.J., Kapsetaki, M. y Margolskee, R.F,				
AUTORES	The exon structure of the mouse Sc1 gene is very similar to the mouse Sparc gene				
TÍTULO	Genome Res. 6 (11), 1077-1083 (1996)				
PUBLICACIÓN	97092869				
MEDLINE	8938431				
PUBMED	2 (bases 1 a 2734)				
REFERENCIA	McKinnon, P.J				
AUTORES	Remisión directa.				
TÍTULO	Remitido (22-JUL-1996) .Gentics, St. Jude Children's Research Hospital, 332 North Lauderdale, Memphis, TN 38101, USA				
PUBLICACIÓN					
CARACTERÍSTICAS	Localización/Cualificadores				
fuelle	1..2734 /organismo="Mus musculus" /db_xref="taxón:10090" /cromosoma="5" /mapa="5E4"				
gen	1..2734 /gen="Sc1" 322..2274 /gen="Sc1"				
CDS	/nota="Similar a Sparc de Mus musculus y a hevina humana; Glucoproteína secretada" /codon_inicio=1 /producto="proteína asociada a matriz extracelular" /id_proteína="AAB08451.1" /db_xref="GI:1546923" /traducción="MKAVLLLLCALGTAVAIFTSTFLSDHNPPTATLVTPEDATVFIAGVEATADIENHPNDKAEPKPSALNSEETHEQSTEQDKTYSFEVDLKDEEDGDGDLSDPTTEGTLTLDLQEGTSEPPQKSLPENGDFPATVSTSYVDPNQRANITKGKESQEQPVSDSHQQPNESKQTQDLKAESSQTQDPDIPNEEEEEEEEEEEEEEPEDIGAPSDNQEEGKEPLEEQPTSKWEGNREQDDTLESSQPTQISKTEKHQSEQGNQGOESDSEAEGEDKASGSKEHIPTEQQDQEGKAGLEAIGNQKDTDEKAVSTEPTDAAVVPRSGGAGDNGGDDSKHGAGDOXFIQSQEFLEAERMHSLSYILKYGGGEETTTGSESNQREAADNQEAKKAESSPNAEPSPDEGNSRHSAGSCTNFQCKRGHICKTDPQGGKPHCVCPQDPETCFPAKILDOACGTDNOTYASSCHLFATKCRLEGTKKGBQLDLYFGACKSIPACTDFFVAAQ				

FFLMRDNLKNIIMQLYEPNPKHGGYLNEKQSKVKKIYLDEKRLLAGDHPIELLLRD
 FKKNYHMYVYPVHWQFNELDQHPADRIITHSELAPLRASLVPMERCITRFFEECDPNK
 DKHITLKEWGHCFGIKEEDIDENLLF" SEC ID N°: 9

RECuento DE BASES 882 a 641 c 663 g 548 t
 ORIGEN

```

1  cagcacggag ggagcgagat ccaggaatct gcaacagaaa ccatgcagac ctgaaacacc
61 ctgtggtgcc aacctccaaa ttctcatctg tcacttcaga cctgactgg ctgacagagc
121 agcagaattt caactccaat aaacgtgaat gtgcttctag gcaaagcagc caagctgacg
181 agggaggggg gtggaagagc tagctoctct tgggcatttg tcaaaactttt aactcctggc
241 tgtgtgcaag gaggggactc aacttcgggt tcaagctacc aaggctctgg atccagccac
301 ctctccgcag atctagccag catgaaggct gtgcttctcc tctgtgccc cttgggaacc
361 gctgtggcaa tcccgacaag tacaagggtt ctctctgacc actccaacc aactactgca
421 acactggtga caccggaaga cgctacagtc ccatgtccg gggttgaagc tacagcagac
481 atagaaaaacc atoccaatga oaaggctgaa aaaccttcag cacttaattc agaagaggaa
541 actcatgaac agtcaacaga gcaggacaaa acctacagct tcgagggtga cctgaaggat
601 gaggaggatg gagatgggga ttaagtgtg gatccaaagg aaggaaact aacactggat
661 ctacaagaag gtacaagtga gctcaacag aaaagtctcc cggaaaacgg ggatttcccc
721 gcgacogtgt ccacttccta tgtggatcct aaccaacggg caaacatcac aaagggaag
781 gagagtcaag agcaacctgt aagtgactca caccagcaac cgaatgaaag cagcaagcaa
841 acccaagact taaaggctga agaaagccag acacaagatc cagacattcc caatgaagaa
901 gaggaagaag aagaggaaga agaagaggaa gaagaggaa agccggaaga cattggtgcc
961 cccagtgata accaagagga gggaaaagaa cctctggagg agcagcctac oagcaagtgg
1021 gaaggaaaca gagagcaatc tcatgacacc ttagaagagt ccagtccagc cactcagata
1081 agcaagacag agaagcatca atctgagcaa ggaaaccaag ggcaggagag tgactctgag
1141 gccaaaggag aggacaaggc ttcaggcagc aaggaaacaca ttccacatac agagcagcag
1201 gaccaagaag ggaagctgg ccttgaagct attggcaacc agaaggacac tgatgagaag
1261 gccgtttcca cagaacctac cgatgctgcc gtggtgccta ggagtcacgg aggagctggt
1321 gataacgggg gccgggatga ctctaagcat ggtgcaggcg atgactactt catccccagc
1381 caggaattcc tagagccdga aaggatgcat tccctctoct attacctcaa atatggcggg
1441 ggcgaggaga caacgactgg cgagagttag aaccagaggg aggtgcaga caaccaagag
1501 gccaaagaaag ctgagagctc accaaatgct gaaccttcag atgagggcaa ctcaagggag
1561 cacagtgtcg gttcttgca gaaacttcaa tgtaaaaggg gacacatttg caaaaaccgat
1621 ccacaaggga aacctactg tgtttgcaa gatccagaga cttgtcccc tgcaaaaatc
1681 ctagatcagg cttgtggcac tgacaaccaa acctacgcaa gctcctgtca cctgtttgct
1741 accaagtga ggtggggg gaocaaaaag ggacaccaac tgcagctgga ttacttcgga
1801 gcttgcaaat ctattcctgc ttgtacggac tttgaagtgg ctgagttcc cctgcggatg
1861 agagactggc tcaaaaaat cctcatgcag ctttatgaac caaatccaa acatggcggc
1921 tatctcaatg aaaagcaag aagcaagtc aaaaaattt acctggatga gaagagactc
1981 ttggtggag accatccat tgaaottoto ttgagggact ttaagaâaaa ctaccacatg
2041 tatgtgtatc ctgtgcactg gcagtttaat gaactggacc agcatcctgc agacaggatc
2101 ttgacacact ctgaacttgc tccctgcga gcttcctggg tgcccatgga acactgcata
2161 actcgcttct ttgaggagtg tgaccccaac aaggataago acatcacctt gaaggaaagg
2221 ggccactgct ttggaattaa agaggaggat atagatgaaa acctootctt ttgaattaa
2281 atttgagaga atcggaactt tccatccacc toacctgtt taacogcttc agaaatacga
2341 gcagccatga cactatacat tcatatgtag oaaaacattt gtttggcatg tgagagaaga
2401 caatggtagt aattacttct tggatgata tatatgagcc aggcactta tattaacta
2461 ggaatgaaa otthaaatt aagtagagtc aatgtctata aaagactgtc ctgtctgggg
2521 acagttagcc accatggcaa tgtcactctg tgcactctgc ttataattg ataattataa
2581 actattaaaa aaacaatgtt catattgtcc ataatacctt atgcatgctg aggaagtga
2641 atactgctct tttgagataa atatgcctcc ttttcagtgt cttggatgtc ctaataaaaa
2701 atctataaaa cccccaaaa aaaaaaaa aaaa SEC ID N°:26

```

179

```

61  caaggctcca gtcacgcag cagttgtcca tgcgtagtt tccactttcc tgtatggcg
121 ggctgggttag gattccactt tcccccaggt gcttagccca gggccagaca aaaagtagtt
181 gcttaagaaa tacttggtga aggaataaat taatgaatga atttgtgctt acagcggctg
241 gatggcagac aaacctatga ttataggaac atcaggatct catttggaac agattacgga
301 tgctgcattg tggaaottgg aaaaacaaga gttottggac aggtttcctg tgaacttgtg
361 tctccaaaac tcaatogggc aacagaaggt attctttttt taaccttgaa ctctctcaga
421 tggccgctcc agctttcgaa cctggcaggc agtcagatct cttgggtgaag ttgaatcgac
481 tcatggaaag atgtctaaga aattcgaagt gtatagacac tgagtctctc tgtgttggtg
541 ctggtgaaaa ggtttggcaa atacgtgtag acctacattt attaatcat gatggaaata
601 ttattgatgc tgcacagcatt gctgcaatcg tggccttatg tcatttcoga agacctgatg
661 tctctgtcca aggagatgaa gtaacaactgt atacacctga agagcgtgat cctgtacctt
721 taagtatcca ccacatgccc atttgtgtca gttttgcctt ttccagcaa ggaacatatt
781 tattggtgga tcccaatgaa cgagaagaac gtgtgatgga tggcttgotg gtgattgcca
841 tgaacaaaaca tcgagagatt tgtactatcc agtcagtggt tgggataatg ctactaaaag
901 atcaagttct gagatgcagt aaaatcgotg gtgtgaaagt agcagaaatt acagagotaa
961 tattgaaago tttggagaat gacaaaaaag taaggaaaga aggtggaaag tttggttttg
1021 cagagtcctat agcaaatcaa aggatcacag catttaaat ggaaaaggcc cotattgata
1081 cctcggatgt agaagaaaaa gcagaagaaa tcattgotga agcagaacct ccttcagaag
1141 ttgtttctac acctgtgota tggactcotg gaactgcca aattggagag ggagtagaaa
1201 actcctgggg tgatcttgaa gactctgaga aggaagatga tgaaggcggg ggtgatcaag
1261 ctatcattct tgatggtata aaaatggaca ctggagtaga agtctctgat attggaagcc
1321 aagatgctcc cataatacto tcagatagtg aagaagaaga aatgatcatt ttggaaccag
1381 acaagaatcc aaagaaaata agaacacaga ccaccagtgc aaaacaagaa aaagcaccaa
1441 gtaaaaagcc agtgaaaaga agaaaaaaga agagagctgc caattaaago taacagttgt
1501 atatctgtat atataactat taaaagggat atttattcca tt SEC ID N°: 27

```




[PubMed](#) [Nucleótido](#) [Proteína](#) [Genoma](#) [Estructura](#) [PMC](#) [Taxonomía](#) [OMIM](#) [Boo](#)

Buscar para

[Vista previa/índice](#) [Historia](#) [Portapapeles](#) [Detalles](#)

1: NM_004684. Homo sapiens SPARCL1 [gi:21359870] [Links](#)

LOCUS	SPARCL1 . 2808 pb ARNm lineal PRI 23-DIC-2002
DEFINICIÓN	Similar 1 a SPARC de Homo sapiens (mast9, hevina) (SPARCL1), ARNm.
REFERENCIA	NM_004684
VERSIÓN	NM_004684.2 GI: 2135987.0
PAL. CLAVE	.
FUENTE	Homo sapiens (humano)
ORGANISMO	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCIA	1 (bases 1 a 2808)
AUTORES	Girard, J.P. y Springer, T.A.
TÍTULO	Cloning from purified high endothelial venule cells of hevina, a close relative of the antiadhesive extracellular matrix protein SPARC
PUBLICACIÓN	Immunity 2 (1), 113-123- (1993)
MEDLINE	95323677
PUBMED	7600298
REFERENCIA	2
AUTORES	Schraml, P., Shipman, R. y Ludwig, C.U.
TÍTULO	An alternative PCR-based method for the direct isolation of cDNA ends (DICE)
PUBLICACIÓN	No publicado
COMENTARIO	REFSEQ PROVISIONAL: Este registro no se ha sometido a una revisión final por el NCBI. La referencia de secuencia procedía de X86693.1 El 9 de junio de 2002 esta versión de secuencia reemplazó a gi: 4758521.
CARACTERÍSTICAS	Localización/Cualificadores
fuelle	1..2808 /ORGANISMO="Homo sapiens" /db_xref="taxón:9606"
.	/cromosoma="4" /mapa="4q22.1" /tipo_tejido="pulmón" /lib_clon="tramo 5' de ADNc en lambda GT10 de pulmón fetal humano (Clontech)
gen	1..2808 /gen="SPARCL1" /nota="sinónimos:HEVINA, MAST9" /db_xref="LocusID:8404" /db_xref="MIM:606041"
variación	complemento (139) /alelo="T" /alelo="A" /db_xref="dbSNP:2615483"
variación	161 /gen="SPARCL1" /alelo="G" /alelo="A"

/db_xref="dbSNP:1049539"
variación 291
 /gen="SPARCL1"
 /alelo="G"
 /alelo="A"
 /db_xref="dbSNP:1130639"
variación 291
 /gen="SPARCL1"
CDS 323..2317
 /gen="SPARCL1"
 /nota="mast9; hevina"
 /codon_inicio=1
 /producto=similar 1 a SPARC"
 /id_proteina="NP_004675.2"
 /db_xref="GI:21359871"
 /db_xref="LocusID:8404°
 /db_xref="MIM:606041"
 /traducción "MKTGPFLLCLLGTAAAIPTNARLLSDHRSKPTAETVAPDNTAIPS
 LWAEAEENEKETAVSTEDDSHHKAEKSSVLKSKEESHEQSAEQGKSSSOELGLKDQED
 SDGHLNVNLEYAPTEGTLDIKEDMIEPQEKLSENTDFLAPGVSSFTDSNQQESITKR
 EENQEQPRNYSHHQLNRSSKHSQGLRDQGNQEQDPNISNGEEEEKEKPGEVGTHNDNQ
 ERKTELPREHANSKQEDNTQSDDILEESDQPTQVSKMQEDEFDQGNQEQEDNSNAEM
 EENASNVNKHITQETEQSQEGKTGLEAISNHKETEETVSEALLMEPTDDGNTTPRN
 HGVDDDDGDDGDDGDTGPRHSASDDYFIPSOAELEAERAQSIAYHLKIEEQREKVHE
 NENIGTTEPGEHQEAKKAENSSNEEBTSSEGNMRVHAVDSMSFQCKRGHICKADQGG
 RPHCVQDQDVTCPPTKPLDQVCGTDNQTYASSCHLFATKCRLEGTKKGHQLQLDYFGA
 CKSIPTCTDFEVIQFPLMRDWLKNILMQLYEANSEHAGYLNKQKQNKVKKIYLDKRL
 LLAGDHPIDLLLRDFKKNYHMYVYPVHWQFSELDQHPMDRVLTHSELAPLRASIVPME
 HCIITREFECDPNKDKHITLKEWGHCFGIKEEDIDENLLF" SEC ID N°: 11
Misc feature 1616..1687
 /gen="SPARCL1"
 /nota="FOLN; Región:similar a dominio N-terminal de folistatina"
 /db_xref="CDD:smart00274"
misc feature 1688..1849
 /gen="SPARCL1"
 /nota="kazal; Región: dominio inhibidor de serin proteasa de tipo
 kazal. Normalmente indicativo de inhidores de serin proteasa.
 Sin embargo, los dominios de tipo kazal también se observan en la
 parte extracelular de agrinas, que no se sabe que sean inhibidores
 de proteasa. Los dominios kazal normalmente aparecen en
 disposiciones en tandem. Plegamiento alfa+beta pequeño que contiene
 tres disulfuros. El alineamiento también incluye un solo dominio de
 transportadores en la familia OATP/PGT"
 /db_xref="CDD:pfam00050"
misc feature 1697..1849
 /gen="SPARCL1"
 /nota="KAZAL; Región: Inhibidores de serin proteasa de tipo kazal"
 /db_xref="CDD:smart00280"
variación 336
 /gen="SPARCL1"
variación 433
 /gen="SPARCL1"
 /alelo="T"
 /alelo="C"
 /db_xref="dbSNP:8342"
variación 433
 /gen="SPARCL1"
variación 458
 /gen="SPARCL1"

variación 468
 /gen="SPARCL1"
 /alelo="C"
 /alelo="A"
 /db_xref=ndbSNP:13051"
variación 468
 /gen="SPARCL1"
variación 638
 /gen="SPARCL1"
 /alelo="G"
 /alelo="C"
 /db_xref="dbSNP:1049544"
variación 638
 /gen="SPARCL1"
variación 668
 /gen="SPARCL1"
variación 702
 /gen="SPARCL1"
variación 1577
 /gen="SPARCL1"
 /alelo="G"
 /alelo="A"
 /db_xref="dbSNP:3189727"
variación 1957
 /gen="SPARCL1"
variación 2047
 /gen="SPARCL1"
 /alelo="T"
 /alelo="C"
 /db_xref="dbSNP:9933"
variación 2047
 /gen="SPARCL1"
variación 2327
 /gen="SPARCL1"
 /alelo="T"
 /alelo="A"
 /db_xref="dbSNP:10626"
 RECuento DE BASES 996 a 554 c 606 g 652 t
 ORIGEN

1 cggcattgaga ggcagcctg ccagggaat ccagggaatc gcaacaaaa cgatgacagt
 61 ctgaaatact ctctgggtgc aacctccaaa ttctcgtctg tcaacttcaga cccccactag
 121 ttgacagagc agcagaatat caactccagt agacttgaat gtgcctcttg gcaaaagaagc
 181 agagctaacg aggaagggga tttaaagagt tttctctggg tggttgctca acttttattc
 241 cctgtctgtg tgcagagggg attcaacttc aattttctgc agtggctctg ggtccagccc
 301 cttacttaaa gatctggaaa gcatgaagac tgggcctttt ttctatgtc tcttgggaaac
 361 tgcagctgca atcccgapaa atgcaagatt attatctgat oattccaaac caactgctga
 421 aacggtagca cctgacaaca ctgcaatccc cagttttatgg gctgaagctg aagaaaaatga
 481 aaaagaaaca gcagtatcca cagaagacga ttcccaccat aaggctgaaa aatcatcagt
 541 actaaagtca aaagaggaaa gccatgaaca gtcagcagaa caggccaaga gttctagcca
 601 agagctggga ttgaaggatc aagaggacag tgatgggtcac ttaagtgtga atttgagta
 661 tgcaccaact gaaggtaact tggacataaa agaagatatg attgagctc aggagaaaaa
 721 actctcagag aacactgatt ttttggtccc tgggtgtagt tcttcacag attctaacca
 781 acaagaaagt atcacaaa gaagaggaaa ccaagaacaa cctagaatt attcacatca
 841 tcagttgaac agtagoagta aacatagcca aggcctaagg gatcaaggaa accaaagagca
 901 ggatccaaat atttccaatg gagaagagga agaagaaaa gagccaggtg aagttggtac
 961 ccacaatgat aaccaagaaa gaaagacaga attgoccagg gagcatgcta acagcaagca
 1021 ggaggaagac aatacccaat ctgatgatat tttggaagag tctgatcaac caactcaagt
 1081 aagcaagatg caggaggatg aatttgatca gggtaaccaa gaacaagaag ataactcaca
 1141 tgcagaaatg gaagaggaaa atgcatcgaa cgtcaataag cacattcaag aaactgaatg
 1201 gcagagtcac gagggtaaaa ctggcctaga agctatcagc aaccacaaag agacagaaga

```

1261 aaagactggt tctgaggctc tgcctoatgga acctactgat gatggtaata ccacgccag
1321 aaatcatgga gttgatgatg atggcgatga tgatggcgat gatggcgga ctgatggccc
1381 caggcacagt gcaagtgatg actacttcac cccaagccag gcctttctgg aggccgagag
1441 agctcaatcc attgocctac acctcaaaat tgaggagcaa agagaaaaag tacatgaaaa
1501 tgaaaatata ggtaccactg agcctggaga gcaaccaagag gccaaagaaag cagagaactc
1561 atcaaatgag gaggaacgt caagtgaagg caacatgagg gtgcatgctg tggattcttg
1621 catgagcttc cagtgtaaaa gaggcacacat ctgtaaggca gaccaacagg gaaaaacctc
1681 ctgtgtctgc caggatccag tgaacttgct tccaacaaaa ccccttgatc aagtttgtgg
1741 cactgacaat cagacctatg ctagtctctg tcatctattc gctactaaat gcagactgga
1801 ggggaccaa aaggggcac aactccagct ggattathtt ggagcctgca aatctattcc
1861 tacttgtaag gactttgaag tgattcagtt tcctctacgg atgagagact ggctcaagaa
1921 tatcctcatg cagctttatg aagccaactc tgaacatgct ggttatctaa atgagaagca
1981 gagaaataaa gtcaagaaaa tttaacctga tgaaaagagg cttttggctg gggacatcc
2041 cattgatctt ctcttaaggg actttaagaa aaactaccac atgtatgtgt atcctgtgca
2101 ctggcagttt agtgaacttg accaacaccc tatggataga gtcttgacac attctgaact
2161 tgctcctctg cgagcatctc tgggtgccat ggaacactgc ataaccogtt tctttgagga
2221 gtgtgacccc aacaaggata agcacatcac cotgaaggag tggggccact gctttggaat
2281 taaagaagag gacatagatg aaaatctcct gttttgaacg aagattttaa agaactcaac
2341 tttccagcat cctcctctgt tctaaccaat tcagaaatat atgcagctgt gatacttgta
2401 gatttatatt tagcaaaatg ttagcatgta tgacaagaca atgagagtaa ttgcttgaca
2461 acaacctatg caccagggtat ttaacattaa ctttggaaac aaaaatgtac aattaagtaa
2521 agtcaacata tgcaaaatac tgtacattgt gaacagaagt ttaattcata gtaatttoac
2581 tctctgcatt gacttatgag ataattaatg attaaactat taatgataaa aataatgcac
2641 ttgtattgtt cataatatca tgtgcaactc aagaaaatgg aatgctactc ttttgtggtt
2701 taogtgtatt attttcaata tcttaatacc ctaataaaga gtccataaaa atccaaaaaa
2761 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa SEC ID N°: 28

```


LOCUS	HSHEVIN	2645	pb	ARNm	lineal	PRI 31-MAR-1995
DEFINICIÓN	ARNm de Homo sapiens para vénula endotelial alta.					
REFERENCIA	X82157					
VERSIÓN	X82157.1 GI5758065					
PAL. CLAVE	Gen de hevina					
FUENTE	Homo sapiens (humano)					
ORGANISMO	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCIA	1					
AUTORES	Girard, J.P. y Springer, T.A.					
TÍTULO	Cloning from purified high endothelial venule cells of hevina, a close relative of the antiadhesive extracellular matrix protein SPARC					
PUBLICACIÓN	Immunity 2 (1), 113-123 (1995)					
MEDLINE	95323677					
REFERENCIA	2 (bases 1 a 2645)					
AUTORES	Girard, J.					
TÍTULO	Remisión directa					
PUBLICACIÓN	Remitido (10-OCT-1994) J. Girard, Center for Blood Research and Dept. of Pathology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA					
CARACTERÍSTICAS	Localización/Cualificadores					
fuelle	1..2645					
	/ORGANISMO="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clon="HEC25"					
	/línea celular="células endoteliales altas purificadas"					
	/tipo celular="células endoteliales altas"					
	/tipo tejido="amígdala"					
gen	1..2645					
	/gen="hevina"					
CDS	198..2192					
	/gen="hevina"					
	/codon_inicio=1					
	/id_proteina="CAA57650.1"					
	/db_xref="GI:758066"					
	/db_xref="SPTREMBL:Q14515"					
	/traducción ="MKTGLFFLCILGTAAPITNARLLSDHSHKPTAETVAPDNTAIPSLRAEDSENEKETAVSTEDDSHKAEEKSSVLSKEESHEQSAEQGKSSSQELGLKDQEDSGDLSTVNLLEYAPSEGTLOIKEDMSEPQEKLSENTDFLAPGVSSFTDSNQQESITKEEENQEQPRNYSHHQLNRSSSKHSQGLRDQGNQEQDPNISNGEEEEKEPGEVGTTHNDQERKTELPREHANSKQEDNTQSDDIILEESDQPTQVSKMQEDEFDQGNQEQEDNSNAEMEEENASNVNKHIEQETWQSQEQGKTGLEAISNHKETEETVSEALLMEPTDDGNTTPRNHGVDGDDGDDGDDGCGTDGPRHSASDDYFIPSQAFLEAERAQSIAYHLKIEEQREKVHENENIGTTEPGEHQEAKKAENSNEEETSSEGNMRVHVDSCMSFQCKRGHICKADQGGKPHCVGCDPVTCPPTKPLDQVCGTDNQTASSCHLEFATKCRLEGTKKRGHQLDLYFG					

CKSIPTCTDFEVIOFFFLMRDNLKNILMQLYEANSEHAGYLNKQNRNVKKIYLDEKR
LLAGDHPIDLLLRDFKKNYHMYVYPVHWQFSELDQHEMDRVLTHSELAPLRASLVFME
RCITRFFEECDPNKDKHITLKEWGHCFGIKEEDIDENLLF" SEC ID N°: 12

198..260

Péptido señal /gen="hevina"

Señal de poliA 2608..2613

/gen="hevina"

RECuento DE BASES 923 a 516 c 580 g 626 t
ORIGEN

```

1 gagcagcaga atttcaactc cagtagaactt gaatatgcct ctgggcacaa aagcagagct
61 aacgaggaaa gggattttaa gagtttttct tgggtgtttg tcaaaactttt attcctgtc
121 tgtgtgcaga ggggattcaa cttcaatttt tctgcagtgg ctctgagtcc agcccttac
181 ttaaagatct ggaagcattg aagactgggc ttttttctct atgtctcttg ggaattgac
241 ctgcaatccc gacaaatgca agattattat ctgacattc caaaccaact gctgaacgg
301 tagcaaccga caacactgca atcccagtt taagggtcga agatgaagaa aatgaaaaag
361 aaacagcagt atcacagaa gacgattccc accataaggc tgaataatca tcagtactaa
421 agtcaaaaaga ggaagccat gaacagtcag cagaacaggg caagagttct agccaagagc
481 tgggattgaa ggatcaagag gacagtgatg gtgacttaag tgtgaatttg gagtatgcac
541 catctgaagg tacattggac ataaagaag atatgagtga gcctcaggag aaaaaactct
601 cagagaacac tgattttttg gctcctgggt ttagttcctt cacagattct aaccaacaag
661 aaagtatcac aaagagagag. gaaaaccaag acaaacctag aattatttca catcatagt
721 tgaacaggag cagtaaacat agccaaggcc taagggtatca aggaaacca ggcagagtc
781 caaatatttc caatggagaa gaggaagaag aaaaagagcc aggtgaagtt ggtacccaca
841 atgataacca agaagaaga acagaattgc ccaggagaca tgctaacagc aagcaggagg
901 aagacaatac ccaatctgat gatatttttg aagagtctga tcaaccaact caagtaagca
961 agatgcagga ggatgaattt gatcagggtt accaagaaca agaagataac tccaatgcag
1021 aaatggaaga ggaatgca tcgaacgtca ataagcacat tcaagaaact gaatggcaga
1081 gtcagagggg taaaactggc ctagaagcta tcagcaacca caaagagaca gaagaaaaga
1141 ctgtttctga ggctctgctc atggaacctc ctgatgatgg taataccacy ccagaaato
1201 atggagttag tgatgatggc gatgatgatg gcgatgatgg ccgcaactgat ggcccagggc
1261 acagtgcagg tgatgactac ttcatcccaa gccaggccctt tctggaggcc gagagagctc
1321 aatccattgc ctatcacctc aaaaattgag agcaaaagaa aaaagtacat gaaaatgaaa
1381 atataggtac cactgagcct ggagagcacc aagaggccaa gaaagcagag aactcatcaa
1441 atgaggagga aacgtcaagt gaaggcaaca tgagggtgca tgctgtggat tcttgcatga
1501 gottccagtg taaaagagga cacatctgta aggcagacca acagggaaaa cctcactgtg
1561 tctgccagga tccagtgaat tgtcctccaa caaaacccct tgatcaagtt tgtggcaetg
1621 acaatcagac ctatgctagt tctgtctac tattogctac taaatgcaga ctggagggga
1681 ccaaaaaggg gcatcaactc cagctggatt attttggagc ctgcaaatct attcctactt
1741 gtacggactt tgaagtgatt cagtttctct tacggatgag agactggctc aagaatatcc
1801 tcatgcagct ttatgaagcc aactatgaac acgctggtta tctaaatgag aagcagagaa
1861 ataaagtcaa gaaaatttac ctggatgaaa agaggctttt ggctggggac catccattg
1921 aacttctctt aagggaactt aagaaaaact accacatgta tgtgtatcct gtgcaactggc
1981 agtttagtga acttgaccaa caccctatgg atagagtctt gacacattct gaacttgctc
2041 ctctgcgagc atctctggtg cccatggaac actgcataac ccgtttcttt gaggagtgtg
2101 accccaacaa ggataagcac atcacctga aggagtgggg ccactgcttt ggaattaaag
2161 aagaggacat agatgaaaat ctcttgtttt gaacgaagat ttaaagaac tcaacttttc
2221 agcatcctcc totgttctaa ccacttcaga aatatatgca gctgtgatac ttgtagattt
2281 atatttagca aaatgttagc atgtatgaca agacaatgag agtaattgct tgacaaccaac
2341 ctatgcacca ggtatttaac attaaacttg gaaacaaaaa tgtacaatta agtaaaagca
2401 acatagcaa aatactgtac attgtgaaca gaagttaaat tcatagtaat ttcactctct
2461 gcattgactt atgagataat taatgattaa actattaatg ataaaaataa tgcatttgta
2521 ttgttcataa tatcatgtgc acttcaagaa aatggaatgc tactcttttg tggtttacgt
2581 gtattatttt caatatctta ataccctaact aaagagtcca taaaaatcca aaaaaaaaaa
2641 aaaaa SEC ID N°: 29

```




[PubMed](#) [Nucleótido](#) [Proteína](#) [Genoma](#) [Estructura](#) [PMC](#) [Taxonomía](#) [OMIM](#) [Boo](#)

Buscar para

☐ 1:X86693. H sapiens ARNm fo...[gi:809026] Links

LOCUS	HSMAS9	2808 pb	ARNm	lineal	PRI 08-SEP-1995
DEFINICIÓN	ARNm para proteína similar a hevina de Homo sapiens.				
REFERENCIA	X86693				
VERSIÓN	X86693.1 GI:809026				
PAL. CLAVE	Proteína similar a hevina; gen MAST9.				
FUENTE	Homo sapiens (humano)				
ORGANISMO	<u>Mono sapiens</u>				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCIA	1				
AUTORES	Schraral, P., Shipman, R. y Ludwig, C.U.				
TÍTULO	An alternative PCR-based method for the direct isolation of cDNA ends (DICE)				
PUBLICACIÓN	No publicado				
REFERENCIA -	2 (bases 1 a 2808)				
AUTORES	Schraml, F.H.				
TÍTULO	Remisión directa				
PUBLICACIÓN	Remitido (26-ABR-1995) P.H. Schraml, Zentrum fuer Lehre und Forschung (ZLF), Molecular Oncology, Laboratory 405, Kantonsspital Basel, Hebelstrasse 20, CH-4031 Basel, SWITZERLAND				
COMENTARIO	Secuencias relacionadas X82157 y S73397.				
CARACTERÍSTICAS	Localización/Cualificadores				
fuelle	1..2808 /organismo="Homo sapiens" /db_xref="taxón:9606" /tipo_tejido="pulmón" /lib_clon="tramo 5' de ADNc en lambda gt10 de pulmón fetal humano (Clontech)"				
<u>gen</u>	1..2808 /gen="MAST-9"				
<u>ARNm</u>	<1..>2808 /gen="MAST-9"				
<u>variación</u>	291 /gen="MAST-9"				
<u>gen</u>	323..2317 /gen="MAST-9"				
<u>CDS</u>	323..2317 /gen="MAST-9" /codon_inicio=1 /producto="proteína similar a hevina" /id_proteína="CAA60386.1" /db_xref="GI:809027" /db_xref="SPTREMBL:Q14800" /traducción ="MKTGPFTLCLLGTAAAIPTNARLLSDHSKPTAETVAPDNTAIPS LWAEAEENEKETAVSTEDDSHHKAEKSSVLKSKESHEQSAEQKSSSQELGLKDQED SDGHLNVNLEYAPTEGTLDIKEDMIEPQEKLSNTDFLAPGVSSFTOSNQESI TKR EENQEQPRNYSHHQLNRSSKHSQGLRDQGNQEQDPNISNGEESKEKEPGEVGTENDNQ				

ERKTELPREHANSKQEDNTQSDDILEESDQPTQVSKMQEDEFDQGNQEQEDNSNAEM
 EEENASNVNKHIOETEWQSQEGKTGLEAISNHKETEERTVSEALLMEPTDDGNTTFRN
 HGVDDGDGDDGDDGDDGPRHSASDDYFIPSOAFLEAERAQSIAYHLKIEEQREKVHE
 NENIGTTEPGEHQEAKKAENSSNEEETSSEGNMRVHAVDSMSFQCKRGHICKADQOG
 KPHCVCDPVTCPPTKPLDQVCGTDNQTAYSSCHLFATKCRLEGTKKGHQLQLDYFGA
 CKSIFTCCTDFEVIQFPLMRDWLKNILMQLYEANSEHAGYLNKQNRNVKKIYLDEKR
 LLAGDHPIDLLRDFKKNYHMYVYPVHWQFSELDQHPMDRVLTHSELAPLRASLVPME
 HCITRFEECDPNKDKHITLKEWGHCFGIKEEDIDENLLF" SEC ID N°: 13

variación 336
 /gen="MAST 9"
 variación 433
 /gen="MAST 9"
 variación 458
 /gen="MAST 9"
 variación 468
 /gen="MAST 9"
 variación 638
 /gen="MAST 9"
 variación 668
 /gen="MAST 9"
 variación 702
 /gen="MAST 9"
 variación 1957
 /gen="MAST 9"
 variación 2047
 /gen="MAST 9"

RECuento DE BASES 996 a 554 c 606 g 652 t

ORIGEN

1 cggcatgaga ggccagcctg ccagggaat ccaggaatct gcaacaaaaa cgtatgacagt
 61 ctgaaatact ctctggtgcc aaactocaaa ttctcgtctg tcacttcaga cccccactag
 121 ttgacagagc agcagaatat caactocagt agacttgaat gtgcctctgg gcaaaagaagc
 181 agagctaacg aggaaggga ttttaasgagt ttttcttggg tgtttgtcaa acttttattc
 241 cctgtctgtg tgcagagggg attcaacttc aattttctgc agtggctctg ggtccagccc
 301 cttacttaaa gatctggaaa gcatgaagac tgggcctttt ttcttatgtc tcttgggaac
 361 tgcagctgca atccgacaa atgcaagatt attatctgat cattocaaa caactgctga
 421 aacggttagca cctgacaaca ctgcaatccc cagtttatgg gctgaagctg aagaaaatga
 481 aaaagaaca gcagtatcca cagaagacga ttcccaccat aagggtgaaa aatcatcagt
 541 actaaagtca aaagaggaaa gccatgaaca gtcagcagaa cagggcaaga gttctagcca
 601 agagctggga ttgaaggatc aagaggacag tgatggtcac ttaagtgtga atttggagta
 661 tgcaccaact gaaggatcat tggacataaa agaagatatg attgagcctc aggagaaaaa
 721 actctcagag aacactgatt ttttggctcc tgggttagt tccttcacag attctaacca
 781 acaagaagt atcacaaga gagaggaaaa ccaagaacaa cctagaaatt attcacatca
 841 tcagttgaac aggagcagta aacatagcca aggcctaagg gatcaaggaa accaagagca
 901 ggatccaaat atttccaatg gagaagagga agaagaaaaa gagccagggtg aagtgtgtac
 961 ccacaatgat aaccaagaaa gaahgacaga attgcccagg gagcatgcta acagcaagca
 1021 ggaggaaagao aatacccaat ctgatgatatt tttggaagag tctgatcaac caaotcaagt
 1081 aagcaagatg caggaggatg aatttgatca gggtaacaaa gaacaagaag ataactocaa
 1141 tgcagaaatg gaagaggaaa atgcatcgaa cgtcaataag cacattcaag aaactgaatg
 1201 gcagagtcaa gagggtaaaa ctggcctaga agctatcagc aaccacaaag agacagaaga
 1261 aaagactgtt tctgaggctc tgctcatgga acctactgat gatggtaata ccacgcccag
 1321 aaatcatgga gttgatgatg atggcgatga tgatggcgat gatggcgga ctgatggccc
 1381 caggcacagt gcaagtgat actactteat cccaagccag gcttttctgg aggcgagag
 1441 agctcaatoc attgcctatc acctcaaat tgaggagaaa agagaaaaag tacatgaaaa
 1501 tgaaaatata ggtacactg agcctggaga gcaccaagag gccaagaaa cagagaactc
 1561 atcaaatgag gaggaacgt caagtgaagg caacatgag gtgcatgctg tggattcttg
 1621 catgagcttc cagtgtaaaa gaggccacat ctgtaaggca gaccaacagg gaaaacctca
 1681 ctgtgtctgc caggatccag tgacttgtoc tocaacaaa ccccttgatg aagtttgtgg
 1741 cactgacaat cagacctatg ctagtctctg tcatctatcc gctactaaat gcagactgga
 1801 ggggacaaa aaggggcatc aactccagct ggattatttt ggagcctgca aatctattcc
 1861 tacttgtacg gactttgaag tgattcagtt tectctaagg atgagagact ggctcaagaa

```

1921 tatcctcatg cagctttatg aagccaactc tgaacatgct gggtatctaa atgagaagca
1981 gagaaataaa gtcaagaaaa tttaacctga tgaaaagagg cttttggctg gggaccatcc
2041 cattgatctt ctcttaaggg actttaagaa aaactaccac atgtatgtgt atcctgtgca
2101 ctggcagttt agtgaacttg accaacaccc tatggataga gtcttgacac attctgaact
2161 tgctcctotg cgagcatctc tggtgcccat ggaacactgc ataaccggtt tctttgagga
2221 gtgtgacccc aacaaggata agcacatcac cctgaaggag tggggcact gctttggaat
2281 taaagaagag gacatagatg aaaatctctt gttttgaacg aagattttaa agaactcaac
2341 ttccagcat octoctctgt totaaccact tcagaaatat atgcagctgt gatacttgta
2401 gatttatatt tagcaaatg ttagcatgta tgacaagaca atgagagtaa ttgcttgaca
2461 acaacctatg caccaggtat ttaacattaa ctttggaac aaaaatgtac aattaagtaa
2521 agtcaacata tgcaaaatc tgtacattgt gaacagaagt ttaattcata gtaattcac
2581 tctctgcatt gacttatgag ataattaatg attaaactat taatgataaa aataatgcat
2641 ttgtattggt cataatatca tgtgcacttc aagaaaatgg aatgctactc ttttgtggtt
2701 tacgtgtatt attttcaata tottaatacc ctaataaaga gtccataaaa atccaaaaaa
2761 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa SEC ID N°: 30

```

[PubMed](#)
[Nucleótido](#)
[Proteína](#)
[Genoma](#)
[Estructura](#)
[PMC](#)
[Taxonomía](#)
[OMIM](#)
[Boo](#)

Buscar

 para

[Límites](#)
[Vista previa/índice](#)
[Historia](#)
[Portapapeles](#)
[Detalles](#)

1: M61908. *C.cornix japoni*...[gi:213614]

Links

LOCUS	QULQR1 2678 pb ARNm lineal VRT 28-ABR-1993
DEFINICIÓN	Gen QR1 de C. coturnix japónica, cds completo.
REFERENCIA	M61908
VERSIÓN	M61908.1 GI:213614
PAL. CLAVE	.
FUENTE	Coturnix coturnix (codorniz común)
ORGANISMO	Coturnix coturnix Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.
REFERENCIA	1 (bases 1 a 2678)
AUTORES	Guermah, M., Crisanti, P., Laugier, D., Dezelee, P., Bidou, L., Pessac, B. y Calothy, G.
TÍTULO	Transcription of a quail gene expressed in embryonic retinal cells is shut off sharply at hatching
PUBLICACIÓN	Proc. Natl. Acad. Sci; U.S.A. 88 (10), 4503-4507 (1991)
MEDLINE	<u>91239596</u>
PUBMED	<u>2034690</u>
COMENTARIO	Texto fuente original: ARN de Coturnix coturnix. De EMBL entrada CCQR1; fechado el 23-JUN-1991.
CARACTERÍSTICAS	Localización/Cualificadores
fuente	1..2678 /organismo="Coturnix coturnix" /db_xref="taxon:9091"
gen	220..2250 /gen="QR1"
CDS	220..2250 /gen="QR1" /codon_inicio=1 /id_proteína="AAA49499.1" /db_xref="GI:213615= /traducción ="MKTVLLICLLGSAFTTPTDPLNYQFGANGOKTAEKHKYTHSEM PEENTGFVNKGDVLSGHRTIKAEVFPVLDTOQKDEPWASRRQCGQDGEHQTKNSLR SINFLTLSHNPGLASDNQESNSGSSRECHSSEHHQPRRHAKHGNMAGQWALRGES PVDALGLVRENTWKYNKNTVGLDENNNNGSEEEAGEEEDDEWGEETDYRDMK HARGTSHGREYRRQNENSRPSGFEFLRDSSLPVRI TKRHGEKFSMEEESQEKLY KEGKLP,SKKNHNE DQGEKRQSEESKEHFQVVNQKRAVTKRQDK EGSNABEDDNDSGDDGEEDLGNVWREAVYEEERMQSNDQDSITNKQKEIT AGDDSGVYREMODYKGDKIKDVTHSEDNHYHH EFPNSSKQQLQTS SSVESMNSTEHEDVKTGGSYHEESARNSTCKALPDLQRNFHC KHGKVCQADKQKGFSCICQDPACPS TKDYKRVCGTDNKTYDGTGCOLFGTK CQLEGTK MGRQLHLDDYMGACKH IPHCTDYEVNQFPLMRD WLKNILMQYIERDQDTSAFLTEKQRNKVKKIYLN EKR LVS GEHPVELLLHDFEKNYHMYLPYVHWQFYQLDQHPVDRSLT HSE LAPLRASLPMEHCITRFFQECDDQDKLITLKEWCHCF AIKEEDINENLLF" SEC ID N°:14
Péptido señal	220..270 /gen="QR1"
Péptido coin	271..2247 /gen="QR1"

/producto="sin nombre"
 220..2250
 /gen="QR1"
 /codon_inicio=1
 /id_proteina="AAA49500.1"
 /db_xref="GI:213616"
 /traducción ="MKTVLLLCILGSAFTTFTDPLNYQFGAHGQKTAEKHKYTHSEM
 PEEENTGFVNKGDVLSGHRITKAIEVFLDTQKDEPWASRRQGGDGEHQTKNSLRISIN
 FLTLHSNPGLASDNQESNSGSSREQHSSEHHQPRRHRKHGNNMAGOWALRGESFVDAIG
 LVRERNTWKYNKNTVGLDENNGSSEEEAGEEEDDEEWGESETDYRDMKHRARGTSHGRE
 YRRWQENENS RPSGEFLRDSSLPVRI TKRHGEKFSMEEESQEKLYKEGKLP LSKKNHNE
 DQGEKRSSEESKEHFOVVNQKRAVTRKQDKEGSNAEEDDNDSDGDDGEDLGNVWRE
 AVYEEERMQSNDQDSITNKQKEITAGDDSGVYREMQDYKGDKI KDVT HSEDNHYHR
 EPFNSSSKQQLQTSSSVESMNSTEHEDVKTGGSYHEESARNSTGKALPDLCRNFHC
 KRQKVCQADKQKPSICQDPAACPSTKDYKRVCGTDNKT YDGT CQLFGTKCQLEGTK
 MGRQLHLDYMGACKHIPHCTDYEVNQFPLMRDWLKNILMQYYERDQDTSAFLETKQR
 NKVKKIYLNKRLVSGEHPVELLLHDFEKNYHMYLYFVHWQFYQLDQHPVDRSLTHSE
 LAPLRASLVPMEHCITRFFQECDDGQDKLITLKEWCHCFATKEEDINENLLF"
 SEC ID N°: 36

RECUENTO DE BASES 898 a 570 c . 619 g 591 t

ORIGEN

1 gtcagacott cctcctcaga agctcacaga aaaacacgct ttctgaaaga ttccacactc
 61 aatgccaaaa tataccacag gaaaattttg caaggctcac ggatttccag tgcaccactg
 121 gctaaccaag taggagcacg tcttctactg coatgaaagg aaaccttcaa accctaccac
 181 tgagccatta actaccatcc tgtttaagat ctgaaaaaca tgaagactgt attgctcctg
 241 atttgtcttc taggatctgc tttcaccact ccaaccgatc cattgaacta ccaatttggg
 301 gccoatggac agaaaactgc agagaagcat aaatatactc attctgaaat gccagaggaa
 361 gagaacacag ggtttgtaaa caaagtgat gtgctgtctg gccacaggag cataaaagca
 421 gaggtaccgg tactggatac acagaaggat gagccctggg cttccagaga acaaggacaa
 481 ggtgatggg agcatcaaac aaaaaacagc ctgaggagca ttaacttctc tactctgcac
 541 agtaatccag ggttggcttc tgataaccag gaaagcaact ctggcagcag cagggaacag
 601 cacagctctg agcaccacca gcccaggagg cacaggaaac acggcaacat ggctggccag
 661 tgggctctga gaggagaaaag tccagtggat gctcttggtc tggttcgtga gcgcaacaca
 721 tggaaataca ataaaaacac agttggccta gatgaaaaca acaatggaag tgaagaagag
 781 gaagctgggg aggaagaaga tgagggaatgg ggtgaagaaa ctgattacag ggatatgaaa
 841 cacagagccc gtgggacaaag coatggaaga gaatacagaa gatggcaaaa tgaanaacagc
 901 cggccatctg gtgaattott gagagattcc agtctgccag tacgtataac caagagacac
 961 ggtgagaaat tcagcatgga ggaggaaagt caggaaaagc tctacaagga aggaaaactc
 1021 cctctctcaa agaaaaatca taatgaggat caaggtgaaa aaagacaaag tgaagaaagt
 1081 aaagagcatt ttcaagtagt caatcagcgc aaacacagag cagtgcagaa aaggcaggat
 1141 aaggagggca gcaatgctga ggaggatgat aatgatagt gtgatgatgg tgaggaagat
 1201 cttggcaatg tctggaggga agcagtctac gaggaagagg aaagaatgca aagcaatgac
 1261 caggacagta tcaactaaca gcaaaaagag gaaataactg ctggagatga cagtggagt
 1321 tatagggaga tgcaggatta caaagggtgac aaaattaaag atgttactca ctctgaagac
 1381 aatcattacc accatgagcp coctaattcc agcagcaagc acaactgca acaagtagc
 1441 tctgttgaga gcatgaattc aacagagcat gaggatgagg ttaagaccac agggagttca
 1501 tatoatgagg aaagtgcagg gaacagcact ggggaaggctc tcccgatct ttgtagaac
 1561 ttccactgca aaagaggaaa agtctgccaa gcagacaagc aaggaaaacc cagctgtatt
 1621 tgccaagatc ctgctgcttg cccttccacc aaagattata agcgtgtttg tggcactgat
 1681 aataagactt acgatggtag gtgccaactc tttggcacca aatgtcaact tgaagggaca
 1741 aaaatgggac gccagctgca cctggactat atgggtgcct gcaaacacat accccactgt
 1801 actgattacg aagtgaatca gttccctctc cgtatgagag actggctcaa aaacatccta
 1861 atgcaatatt atgaacgtga tcaggatagc tctgcatttc taacggaaaa gcaaggaat
 1921 aagggtcaaaa agatatacct gaatgagaag cgtctctctc ctggtagaca ccacgttag
 1981 cttctctctg atgactttga gaaaaactac cacatgtate tctatcctgt gcaactggcaa
 2041 ttttatcagc ttgaccagca cccagttgac agatcaactga ctcatcaga gctcgtctct
 2101 ttgagagcct ccctcgttcc catggaacac tgcataacc gtttcttcca ggagtgtgat
 2161 ggagaccag acaaaactat caotttgaaa gagtgggtgc actgotttgc gattaaggaa
 2221 gaagacataa atgaaaatct cctgttctga gccacactga gcagaatccc catgcagcgc
 2281 tacagcttgt caaacatgca atgcccattt atgaactgaa ttaacagctc tgtaatttt
 2341 cagggaataag ttggcataag attcttggag gcagaacaag tctctcttgg ataacacaag
 2401 tgcctaattg ttacaaattc attaacagca gtagtgttta agagctctaa gtagctcata
 2461 cttaagagat gtttccctct gcaogtaaca ataactctct agtaagaca ctaacttgat
 2521 gactgagttg ttcaaaaaac ccttccgtag aattatagga tgtggatttt ataatacacc
 2581 gataaaaaact actttgaaat aggttttctt tctctgtcgt ttaactgtcag tagctctctg
 2641 catagaaatg tcaaataaac agatcttgtt ttggtttc SEC ID N°: 31




[PubMed](#) [Nucleótido](#) [Proteína](#) [Genoma](#) [Estructura](#) [PMC](#) [Taxonomía](#) [OMIM](#) [Boo](#)

Buscar para

[Limites](#) [Vista previa/índice](#) [Historia](#) [Portapapeles](#) [Detalles](#)

1: NM_008047. Mus musculus foll... [gi:6679870] [Links](#)

LOCUS	Fst1	2823 pb	ARNm	lineal	ROD 19-SEP-2002
DEFINICIÓN	Similar a folistatina de Mus musculus (Fst1), ARNm.				
REFERENCIA	NM_008047				
VERSIÓN	NM_008047.1 (GI: 6679870)				
PAL. CLAVE	.				
FUENTE	Mus musculus (ratón doméstico)				
ORGANISMO	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCIA	1 (bases 1 a 2823)				
AUTORES	Shibanuma, M., Mashimo, J., Mita, A., Kuroki, T. y Nose, K.				
TÍTULO	Cloning from a mouse osteoblastic cell line of a set of transforming-growth-factor-beta 1-regulated genes, one of which seems to encode a follistatin-related polypeptide				
PUBLICACIÓN	Eur. J. Biochem. 217 (1), 13-19 <1993>				
MEDLINE	94039028				
PUBMED	7901004				
COMENTARIO	REFSEQ PROVISIONAL: Este registro no se ha sometido a una revisión final por el NCBI. La referencia de secuencia procedía de M91380.1				
CARACTERÍSTICAS	Localización/Cualificadores				
fuelle	1..2823 /organismo="Mus musculus" /db_xref="taxon:10090" /cromosoma="16" /mapa="16 27.3 cM" /línea_celular="MC3T3-E1" /tipo_celular="osteoblasto"				
gen	1..2823 /gen="Fst1" /nota="sinónimos: TSC-36, AI316791, AW107808" /db_xref="LocusID:14314" /db_xref="MGD:102793"				
CDS	80..1000 /gen="Fst1" /codon_inicio=1 /producto="similar a folistatina" /id_proteina="NP_032073.1" /db_xref="GI: 6679871" /db_xref="LocusID:14314" /db_xref="MGD:102793" /traducción="MWKRNIALSLVTIALVHGEEPRSKSKICANVFCGAGRECAVTE KGEPTCLCIEQCKPKRPVCGSNGKTYLNHCELHRDACL TGSKI QVDYDGHCKEKKSA SPSASFVVCYQANRDELARRLIQWLEAEITPDGWFSKGSNYSEILDKYFKSFDNGDSH LDSSEFLKFEQNETAINITTYADQENNKLLRSLCVDALIELSDENADWKLSTQEFLLK CLNPSFNPFKKCALEVETIADGAETEVDNRCVCSGHWVCTAMTCDGKNQKGVQTH TEEKTGYVQELQKHQGTAEKTKKVNTEI" SEC ID N°: 15				
misc feature	233..367				


```

/gen="Fst1"
/nota="KAZAL; Región: inhibidores de serin proteasa de tipo
kazol"
/db_xref="CDD:smart00280"
misc feature 239..367
/gen="Fst1"
/nota="kazal; Región: dominio inhibidor de serin proteasa de
tipo kazal. Normalmente indicativo de inhibidores de serin
proteasa. Sin embargo, los dominios de tipo kazal también se
observan en la parte extracelular de agrina, que no se sabe que
sean inhibidores de proteasa. Los dominios kazal a menudo
aparecen en disposiciones en tándem. Pliegue alfa+beta pequeño
que contiene tres disulfuros. El alinmaineto también incluye
un solo dominio de transportadores en la familia de OATP/PGT".
/db_xref="CDD:pfam00050"
misc feature 770..913
/gen="Fst1"
/nota="VWC_out; Región; dominio de tipo C de factor de von
Willebrand (vWF)"
/db_xref="CDD:smart00215"
variación 1362
/gen="Fst1"
/alelo="A"
/alelo="C"
/db_xref="dbSNP:3023433"
RECuento DE BASES 762 a 663 c 697 g 701 t
ORIGEN
1 aagcgacgct cccaccttcg cctctaactc gctgcgcgca cccctgccag tgcctccgg
61 agtcccggac ccgagcagca tgtggaaacg atggctggcg ctctcgctgg tgaccatcgc
121 cotgggtccac ggcgaggagg aacctagaag caaatccaag atctgcgcga atgtgttttg
181 tggagctggo agggaatgtg cgtgcacaga gaagggggag ccoacgtgco totgcattga
241 gcaatgcaaa cctcacaaga ggcctgtgtg tggcagtaat ggcaagaact acctcaacca
301 ctgtgaactt catagagatg cctgcctcac tggatccaag atccagggttg atttatgatg
361 gcactgcaaa gaaaagaagt ctgcgagtc cctgtccagc ccagttgtct gctatcaagg
421 taaccgcat gactccgac ggcgcctcat ccagtggtg gaagctgaga toattccaga
481 tggctggttc tctaaaggca gtaactacag tgagatccta gacaagtaot ttaagagctt
541 tgataatggo gactctcacc tggactccag tgaattcctg aaattcgttg agcagaatga
601 aacagccatc aacatcaca cttatgcaga tcaggagaac aacaaaotgc tcagaagcct
661 ctgtgttgac gccctcattg aaotgtotga tgagaacgct gactggaaac tcagcttcca
721 agagttccto aagtgcctca acccatcctt caacctcctt gagaagaagt gtgcccgtga
781 ggtcgaaacc tatgcagatg gactgtgagc tgaggtggac tgcaatcgt gtgtctgttc
841 ctgtggccac tgggtctgca cagcaatgac ctgtgatgga aagaatcaga aggggggtcca
901 gaccacacac gaggaggaga agacaggata tgtccaggaa ctccagaago accagggcac
961 agcagaaaaa accaagaagg tgaacaccaa agagatctaa gaagaggcac agagcacogt
1021 gtcgggagcb cagcgcctcc tcttcagcgc tgagccagt acacacagag tctgcagcaa
1081 tcaccnaatc actagtattt gcttgtatgg cagcgaatct tatctgttt gttttgcaat
1141 aaaggaaatg aggggtggca gcctagcag ggaaggccac aaccttcaco tgtagggaatg
1201 ctttaagaga aactaaagga cacttgagg cagagaggca ctaaggaaac agcatcgggt
1261 tggcagagga gcagaggcag gtttgaatga agcctttctg gggtcacagc agctgcagg
1321 agaatacagg aaaagcatag agaaacattg aactagccct gctggaggaa gtggggggag
1381 ctttgtaggg aggaaccctg ctgctttgac oottgtcaco aotgtcagoa tgacagacot
1441 gcagcaagtc tgcctctcct ttgggtccca acaatcaoot gaacacacag ccgcccact
1501 agttaactgt gtcctcagcc ttgcatggag tttcctggag gaggtgttta aatgatgcag
1561 acacttatgt actcaagcg catggagact aacpaaattt ttaaaataca tttttctttt
1621 tttttttttt ttgttaacca aagggtctat ttctctgtaa aagacttttt tccaagctga
1681 cttcattcct cagttattac cgttatatta ttgttgtttt ttaattttt attttttgac
1741 tagatattaa gctttttgta ttatttttca ttagtctac tatttcgaga agtgaagggtg
1801 aaggggggtt gggcattttt ccagggtaca gggaaactct taacacaaac agcccatacc
1861 ctgtcacata ttgacccggt tgcagttcgg agcatgcaco ccaaccaga gcttctagaa
1921 aatcagctcc atgccacgaa ggcacaagag gccctcagc agaagccaca ggacaaagca

```

```

1981 tcttcataga cagctgttga gatccaaaoc gttaatttgc ttttgtttct tgtaagaagt
2041 tccaaggatg gacgctcagg ctatcccagc ctgccagcct gctgtgatct gtggctaact
2101 ggcagagtca gcaactgtgg tccttagctg ctctgtttc taggtgtcag ttacttagt
2161 aaactggtaa gaatgaatct tggaaattta taaatggtag tttgtggttt agccaactgg
2221 tccagagggg gctaccttct ccttaggata gatgaatcta ctccataaga aaaaccagcc
2281 aggaatagca tggatgggtt ttgctttggt tgaatgatc ctacgaggtg actgggtatg
2341 aggacttcat ggtcactctg cccaggaaga gacgctgaag gacaactagc agcttcctta
2401 gggatggtac acatgtgtgt gatctctgga gatcagaggt tgccccacac acatgatgat
2461 aaaacttttc agatttagag cggttaaaac tggagatcga atctggattg agaatcagca
2521 ctgggggcag aaactgttat tgaaagtcaa tcctttcttt gagacactcc gaataaacta
2581 tggagatttt cctgcatagg aaagtgtgga atgttgagct attgagatgg gagtggattt
2641 cgtcctaaat agttttttcc tggctctcat tgaacaagac aatttgctct gctagtgtt
2701 ctgtgacctc cctttcaaaa gctctgagcc ccgctcatgc agtccagatt tcatccccct
2761 ctccaagtgc cttggagagc tcacgacagc aatgccatca tcaaaagttt tgotgtggg
2821 aag SEC ID N°: 32

```

Links

```

1  cactctotgt  tgtccaatgg  acacacctgt  cgtgttttga  gcaagcgaga  gatgcagtgg
61  aagtgaaaag  catggttaca  gactcccat  ggcacagta  actctttctga  agtagcggac
121  gcttggttag  cttgacattc  tatgcaaaga  tcataaatgt  ggttcctgca  gatggcacag
181  ttatcaacca  caatatccca  ggccagagg  gctactgcat  tccacttttt  caottcaaag
241  cgcttcttgc  ccgcgcgcgt  gttggtgccg  ctgggggtat  ccacatccat  cgctgcgggc
301  tcacaaagcg  ccgcagacgt  cggcgggcgg  gtgtggcaag  agcgcagggg  cgcgagccgg
361  cgatcagcct  tccggcgac  cgtgccggcg  gagctcgagc  aactcggact  aggggacccg
421  ggccggcccc  caagatgcgg  gcgatcgcg  tgttggcgcg  ggccgcgcgg  gcgtggtgct
481  tcttccaagt  cgagagccgg  cacctggag  cgtctccggc  agggcggggc  cccaaccacg
541  gcaatttctt  agacaatgac  cagtggctga  gcacgctctc  ccagtacgac  cgggacaaat
601  actgqaaccg  ctttcgagac  gatgattatt  tcagaactg  gaattccaac  agcccttttg

```

661 accaagccct ggacccatcc aaggaccctt gctgaaggt aaaatgcagc cotcacaag
 721 tgtgtgtgac ccaggactac cagaccggcc tgtgtgtcag ccgcaagcac ctgtctccoa
 781 ggcaaaaagaa ggggaacgtg gcccaaaaaa actgggttgg accttcgaat ttggtcaagt
 841 gcaagccctg tccctggca cagtccagca tggctgtcgg ctccagatggc cactccatca
 901 catccaagtg caaatggag ttccatgctt gttctactgg caaagccctc gccacccctc
 961 gtgatgggoc ctgtccctgt ctccagagc ctgagccacc aaagcacaaag gcagaaagga
 1021 gtgcctgcac agacaaggag ttgcggaacc ttgcctcccg gctgaaggat tggtttggag
 1081 ctctccacga ggatgcgaac agagtcatca agccaccag ctccaaacaca gcccaaggca
 1141 ggtttgacac tagcatcctg cccatctgca aggaactccct gggctggatg ttcaacaagt
 1201 tggacatgaa ctatgacctc ctgcttgacc ctccagagat caatgccatc taactggata
 1261 agtacgagcc ctgtatcaag cctcttttca actcgtgtga ctccctcaag gatggcaagc
 1321 tttctaaaca tgagtgggco tactgttcc agaagcctgg aggtctccct tgcagaatg
 1381 aaatgaacga aattcagaag ctgagtaagg ggaagccct gttggggggc ttcatatctc
 1441 ggtgtaatga ggagggctat tacaagcca cacagtcca cggcagacg gggcagtgct
 1501 ggtgtgtgga caaatatggg aatgagttgg ctggctccag gaaacagggc gctgtgagct
 1561 gtgaagagga gcaggaaccc tcaggggatt ttggcagtg tgggtccgtg gtccgtgctg
 1621 atgacctaga atatgaacgg gagctgggac caaaggacaa agaggggaag ctgaggggtc
 1681 acaccggagc cgtgacagag gatgatggg atgaggatga tgacaaagag gatgaggtcg
 1741 ggtacatatg gtatgcccac caagaaagag gacacaagtt ttgcaaaaaa ttgcaagtca
 1801 ctctctatcc ctgcatttgt atctaagact ccaaggaacc aaggtctctt ctccattgtt
 1861 gctctctata cccgacctaa ggtttggaag acaactgctt gtccccagag gattctgatt
 1921 ttgcatatgt ttgtatggga gaaaggggtg tgtgttttt tttttgttgt tgtttatttt
 1981 ttggataggg aagtcattgg cttaattaga gctccttcc tttctgtgag atttttccaa
 2041 caagcatgtg atttaccgtg aattctgaca gtgcagggag oooooaccc cttaaatgtc
 2101 aagacccctt tttgattacc cactctgggt gttattacag catggttccc agccttacag
 2161 tgtctaaagt cttctcttgt gtccgttaga tgttgtgaaa aagaaaaaaa caaaaaatac
 2221 accacactgt actttttccc cctgcccccg ttactgcccg tgattattat taaaaattag
 2281 tttttttcac atcattctat ctggcttccct ataaacaaca gccttaattc agtcaagact
 2341 ccctttggga attcatttta ttaaaaattg gtgtctggat acttccctgt acatgcataa
 2401 atatgcattg atgtacagaa agactgtatg tgtgtgcctt gcacacacac ccatacctct
 2461 cagaaaaagt gtttgggtat cttaaaaact cgaaaaaaaa tgataaattt ctccagctgt
 2521 ccagacctgg aacaaaaatt ctggaataag aaatttgtat taaagtcctt ttttgacta
 2581 acagttgggt cttgtagcct gcaggtgag gaagtctctt ctctgtgcat cagcagagtt
 2641 actgaaggo tctgattgag aaaaaaccc cgtctgccta aatcactttt ctccgagaag
 2701 ccattgcgact cccacacgac acgggcagct tcacaagcca tctctttcat ttctgtttga
 2761 agcccttgg ctgcagcaat cctgtctgcc ataggtttct tcttctctta cctactcaag
 2821 ggtttttct aaggcatgca cacatatctc ctgttctctg agagtaacct ggtgttccct
 2881 aaaagaagaa aatttctaat totgaatca atgttttgt tttactccct ttctactgac
 2941 aaatcatgat aagggcacaa aagctgtaca gatttttttt ttaaccact caatccaaa
 3001 tggaggccta caaagaacat cgtaatcaaa catggaagca aaccocgggt ttttaagago
 3061 aaattctgtc ccccccac tcccccaagt gacaagatac taatgaagaa agttcttcac
 3121 catagtgttt gttttaapta aaactcattg agtctagttc caaatttggg agggctcatc
 3181 tctctacatt ccttaggatt tttctcccta tcaagctggc ccagatacaa gtacaaaaca
 3241 gtagtctctg aagttcccat ttcttccagt accagctctat aagctactgt ccgccactga
 3301 ttttcatcta tcaggtgtc ctaatcagaa tcagccacco aagcaagcct ctctggtcca
 3361 catatctatc tcttgcttcc ccccatgaac ttccagctgt ccacacaaaa gccacataaa
 3421 ctcaagcaag aaatatgttc agcaaaaaa tgattatagt ggcagctgac caatacccca
 3481 cccc SEC ID N°: 33

1: NM_022137. Homo sapiens secr... [gi:24475656] [Links](#)

197

péptido señal 255..332
/gen="SMOC1"

péptido coin 333..1556
/gen="SMOC1"
/producto="SMOC1"

misc feature 375..515
/gen="SMOC1"
/nota="kazal; Región: dominio inhibidor de serin proteasa de tipo kazal. Normalmente indicativo de inhibidores de serin proteasa. Sin embargo, los dominios de tipo kazal también se observan en la parte extracelular de agrinas, que no se sabe que sean inhibidores de proteasa. Los dominios kazal normalmente aparecen en disposiciones en tándem. Pliegue alfa+beta pequeño que contiene tres disulfuros. El alineamiento también incluye un solo dominio de transportadores en la familia OATP/PGT"
/db_xref="CDD:pfam00050"

misc feature 378..515
/gen="SMOC1"
/nota="KAZAL; Región: inhibidores de serin proteasa de tipo Kazal"
/db_xref="CDD:smart00280"

misc feature 537..728
/gen="SMOC1"
/nota="tiroglobulina_1; Región: repetición de tipo 1 de Tiroglobulina. Se cree que las repeticiones de tipo 1 de Tiroglobulina están implicadas en el control de la degradación proteolítica. El dominio normalmente contiene seis cisteínas conservadas. Estas forman tres puentes disulfuro. La cisteína 1 se empareja con la 2, la 3 con la 4 y la 5 con la 6."
/db_xref="CDD:pfam00086"

misc feature 597..737
/gen="SMOC1"
/nota="TY; Región: repeticiones de tipo 1 de tiroglobulina"
/db_xref="CDD:smart00211"

misc feature 933..1130
/gene="SMOC1"
/nota="tiroglobulina_1; Región: repetición de tipo 1 de Tiroglobulina. Se cree que las repeticiones de tipo 1 de Tiroglobulina están implicadas en el control de la degradación proteolítica. El dominio normalmente contiene seis cisteínas conservadas. Estas forman tres puentes disulfuro. La cisteína 1 se empareja con la 2, la 3 con la 4 y la 5 con la 6."
/db_xref="CDD:pfam00086"

misc feature 996..1136
/gen="SMOC1"
/nota="TY; Región: repeticiones de tipo 1 de tiroglobulina"
/db_xref="CDD:smart00211"

variación 956
/gen="SMOC1"
/alelo="T"
/alelo="C"
/db_xref="dbSNP:3825739"

variación complemento (2487)
/alelo="T"
/alelo="C"
/db_xref="dbSNP:2273780"

variación 2939
/gen="SMOC1"
/alelo="G"
/alelo="A"

```

/db_xref="dbSNP: 3742912"
variación 2968
/gen="SMOCL"
/alelo="G"
/alelo="A"
/db_xref="dbSNP: 3742913"
variación 2993
/gen="SMOCL"
/alelo="T"
/alelo="C"
/db_xref="dbSNP: 3742914"
señal de poliA 3648..3653
/gen="SMOCL"
BASE COONT 842 a 988 c 1026 g 813 t
ORIGEN

1 gcctgctgcc gcctgggccc cgcggagcgg agctagcgcc gcgcgcagag cacacgctcg
61 cgctccagct cccctcctgc gcggttccat actgtgtccc ctgaccgcag cctctgctcg
121 ccccccgcgc aggaccacgg cccgtccccc gcgcgcgcga gggccccgag cgaagggaag
181 aaggaggagg cgctgtgcgc cccgcggagc ccgcgaaccc cgctcgctgc cggctgcccc
241 gcctggcctg caccatgctg ccgcgcgcgt gcgcgcgcgt gctcaccgcc cactgtctgc
301 tgggtgttgt goagctgtcc cctgtctcgg gccacgcgac cacaggcccc aggtttctaa
361 taagtgcacg tgaccacacg tgcaacctcc actgtctcag gactcaaccc aaacccatct
421 gtgcctctga tggcagggtc tacgagtcga tgtgtgagta ccagcgagcc aagtggcgag
481 acccgacct gggcgttgtg catcgaggta gatgcaaaag tgcctggcag agcaagtgtc
541 gcctggagcg ggctcaagcc ctggagcaag ccaagaagcc tcaggaaagt gtgtttgtcc
601 cagagtgtgg cgaggatggc tcttttacc aggtgcagtg ccatacttac actgggtact
661 gctggtgtgt caccocggat gggaagccca tcagtggcto ttctgtgcag aataaaactc
721 ctgtatgttc aggttcagtc aocgacaagc ccttgagcca ggttaactca ggaaggaaaag
781 atgacgggtc taagcgcaca ccacgatgg agaccagcc ggtgttcgat ggagatgaaa
841 tcacagcccc aactctatgg attaaacact tgggtatcaa ggactccaaa ctgaacaaa
901 ccaacataag aaattcagag aaagtctatt cgtgtgacca ggagaggcag agtgccctgg
961 aagaggccca gcagaatccc cgtgagggtg ttgtcatccc tgaatgtgcc cctgggggac
1021 tctataagcc agtgcaatgc caccagtcca ctggctactg ctggtgtgtg ctggtggaca
1081 cagggcgccc gctgcctggg acctccacac gctacgtgat gccagttgtg gagagcgacg
1141 ocagggccaa gactacagag gcggtatgac ccttcaagga cagggagcta ccaggtctgc
1201 cagaagggaa gaaaatggag tttatcacca gctactgga tgcctcacc actgacatgg
1261 ttccaggccat taactcagca gcgcacactg gaggtgggag gttctcagag ccagaccoca
1321 gccacacctt ggagggagcg gtagtgcact ggtatttcag ccagctggac agcaatagca
1381 gcaacgacat taacaagcgg gagatgaagc ccttcaagcg ctacgtgaag agaaaagcca
1441 agcccaagaa atgtgcccg cgtttcaocg actactgtga cctgaacaaa gacaaggta
1501 tttcactgoc tgagctgaag ggcctgcctg gtgttagcaa agaaggacgc ctctcttaag
1561 gagcagaaaa cccaagggca ggtggagagt ccagggaggc aggatggatc accagacacc
1621 taaccttcag cgttgcccat ggccctgccca catcccggtg aacataagtg gtgcccacca
1681 tgtttgcact ttttaatact cttacttgcg tgttttgttt ttggtttcat tttaaaacac
1741 caatatotaa taaccagtg ggaaaaggaa agggaaagaa gactttatct totctttat
1801 tgttaagttt tggatctgct actgacaaat tttagagggt tttggggggg tgggggaggg
1861 tgttgttggg gcctgagaag aaagagattt atatgctgta tataaatata tatgtaaatt
1921 gtatagttct tttgtacagg cattggcatt gctgtttgtt tatttctctc cctctgctg
1981 ctgtgggtgg tgggcacatc ggacacatag tccagcttcc taaaatccag gactctatcc
2041 tgggcctact aaacttctgt ttggagactg acccttgtgt ataaagacgg gactcctgca
2101 attgtactgc ggaatccacg agttcttttc tgggtggagg actatattgc cccatgcat
2161 tagttgtcaa aattgataag tcaactgggt ctggccttgc tccagggagg ttgggctaag
2221 gagagatgga aaotgcctg ggagagggaag ggagtcacga tcccatgaat agcccacaca
2281 ggtacoggct ctacaggggt cagtgcattc ctgctctccg gaccccaaaa gggccagca
2341 ttggtgggtg caocagtatc ttagtaccc tgggagcaaa ttatocacaa aggatttgca
2401 ttacgtcact cgaaacgttt tcatccatgc ttagcatcta ctctgtataa cgcagagag
2461 gggaggcaaa gaagaaaagg acacacggaa gggcctttaa aaaagtagat atttaatatc
2521 taagcagggg aggggacagg acagaaagcc tgcactgagg ggtgcggtgc caacagggaa
2581 actcttcacc tccctgcaaa cctaccagtg aggtcccgag agacgcagct gtctcagtcg
2641 ccaggggcag attgggtgtg acctctccac tccatcatct cctgctgttg tccatagtcg

```

```

2701 tatcacaggc ctgggtgggt ggggtggggg aagtgtcagt caccttggtg gtaacactaa
2761 agttgttttg ttggtttttt aaaaacccaa tactgaggtt cttoctgttc cctcaagttt
2821 tcttatgggc ttccaggctt taagctaatt ccagaagtaa aactgatctt gggtttccta
2881 ttctgocctc cctagaaggg cagggggtgat aaccagcta cagggaatc cgggccagc
2941 ttccacagg catcacaggc atcttcgcgc gattctaggg tgggtgccc agccttctgg
3001 tctgaggcgc agctccctct gccaggtgc tgtgcctatt caagtggcct tcaggcagag
3061 cagcaagtgg cccttagcgc cccttcctat aagcagctgt ggtggcagtg agggaggttg
3121 ggtagccctg gactggtccc ctctcagat cacccttgca aatctggcct catcttgtat
3181 tccaaaccga catccctaaa agtacctcca ccggttcggg gtctggaagg cgttggcacc
3241 acaagcactg tcctctgtgg aggagcaca ccttctcggg acaggatctg atggggctctt
3301 gggctaaagg aggtccctgc tgtcctggag aaagtccctag aggttatctc aggaatgact
3361 ggtggccctg cccaacgtg gaaaggtggg aaggaagcct tctccatta gccccaatga
3421 gagaactcaa cgtgccggag ctgagtgggc cttgcacgag acactggccc cactttcagg
3481 cctggaggaa gcatgcacac atggagacgg cgcctgctg tagatgtttg gatcttcag
3541 atctcccccag gcatcttgtc tccacagga tcgtgtgtgt aggtggtgt gtgtggtttt
3601 cctttgtgaa ggagagagg aaactatttg tagcttggtt tataaaaaat aaaaaatggg
3661 taaatcttg SEC ID N°: 38

```





PubMed Nucleótido Proteína Genoma Estructura PMC Taxonomía OMIM Boo

Buscar para Ir

Límites Vista previa/índice Historia Portapapeles Detalles

Presentar 20

1: BC004617. Mus musculus, Sim...[gi:13435493] [Links](#)

LOCUS	BC004617 1611 pb ARNm lineal ROD 20-SEP-2002 •
DEFINICIÓN	Mus musculus, Similar a factor de transcripción 19 (SC1), clon transcription factor. 19 (SC1), clon MGC:6711 IMAGE:3585301, ARNm, cds completo.
REFERENCIA	BC004617
VERSIÓN	BC004617.1 GI:13435493
PAL. CLAVE	MGC.
FUENTE.	Mus musculus (ratón doméstico)
ORGANISMO	<u>Mus musculus</u> Eukaryota; Metazoa; Chordata-; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCIA	1 (bases 1 a 1611)
AUTORES	Strausberg, R.
TÍTULO	Remisión directa
PUBLICACIÓN	Remitido (21-MAR-2001)- National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD -20892-2590, USA
OBSERVACIÓN	URL del proyecto NIH-MGC: http://mgc.nci.nih.gov
COMENTARIO	Contacto: Servicio de asistencia de MGC Email: cgapbs-r@mail.nih.gov Obtención de tejidos: Gilbert Smith, Ph.D. Preparación de la biblioteca de ADNc: Life Technologies, Inc. Biblioteca de ADNc organizada por: El consorcio I.M.A.G.E./LLNL Secuenciación de ADN por: Baylor College of Medicine human Genome Sequencing Center Código de centro: BCM-HGSC Página web: http://www.hgsc.bcm.tmc.edu/cdna/ Contacto: amg@bcm.tmc.edu Gunaratne, P.H., García, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A. Distribución del clon: La información de distribución de clones de MGC puede encontrarse a través del consorcio I.M.A.G.E./LLNL en: http://image.llnl.gov Serie: Placa IRAK: 10 Fila: 1 Columna: 11. Este clon se seleccionó para su secuenciación de longitud completa porque pasó los siguientes criterios de selección: predicción de gen GenomeScan, similitud pero no identidad a la proteína.
CARACTERÍSTICAS	Localización/Cualificadores
fuelle	1..1611 /organismo="Mus musculus" /raza="FVB/N" /db_xref="taxón:10090" /clon="MGC:6711 IMAGE:3585301" /tipo_tejido="tumor mamario. Modelo de metalotioneína-TGF alfa ratón virgen de 10 meses de edad. Obtenido mediante biopsia. /lib_clon="NCI_CGAP_Mam1"

CDS

```

/lab_anf="DH10B"
/nota="Vector; pCMV-SPORT6"
360..1151
/codon_inicio=1
/producto=similar a factor de transcripción 19 (SC1)"
/id_proteína="AAH04617.1"
/db_xref="GI:13435494"
/ traducción = "MLPGFQLLRIGGGGGDLYTFHPPSKSGCTYGLGCRADLCDAVAL
RPQQEPGLISGVHAELHAEHQDDWRVSLSDHSSQGITLVNNVALPRGHRLELSDGDL
TFGPQGGAGTSSSEFYEMFQQVRVKPDFAALTVPFSKGEAGGGTQPMLEPPQGAPO
PLSTLSSAFKATLILNSIGSLKLAQPLTFRRGGGRPQGLAIPSQHGEAQVSPAPPT
RNRKSAKKVLAELDDVSPGPLSVLTFPRKRLRVEKAALIASGE" SEC ID N°: 17

```

RECuento DE BASES 393 a 456 c 4.17 g 345 t
ORIGEN

```

1 ggaatccctt cctcagttta cccgcgcgagc gaaaaaaaga aaggattgtc ggaggctccc
61 ggatgttgag cgccttgagg atccggcttc actgagggtc cacaagaagt cttcctggtt
121 ttgcatcgca tcaggacttt actggaattg ctttttaaac ctcatcttga cagcctctgt
181 cttggtgcat atccaggcct ctggogctct gaatctgcac tgottaagtt ggcattaggo
241 acaataacta gctgttgaca tcagagagaa ctgaattgaa gccacaggtg catgacccag
301 gagtaggaga ggaatggac tttgaggtg ggogaagtg tgacagaggg gcaccccca
361 tgcgccttg tttccagctg ctgbgtatag ggggcggcag gggcgggtgac ctctacacct
421 tccacccccc gtccaagtct ggcctgcacct atggattggg ctgcaggggc gacctgtgtg
481 atgtggccct gcggcoccag caggagcccg gctcatctc tggagtccat ggggaattgc
541 atgctgaact ccaaggggac gactggaggg tcagcctgga ggatcacagc agccaaggga
601 ctttgggtcaa taatgtccga cttccaagag gccacagact ggagttgagt gatggtgacc
661 tttcgacctt tggcccccda ggtcaagcag gaaccagcto ctctcagag ttctacttca
721 tgtttcaaca agtccgggtt aaacctcagg actttgcage cataaccgtc cctcgggtota
781 agggagaage tggggggcgt tttcaacctc tgctgcccc ccaaggggca cctcagaggo
841 cactcagcac tctctctctt gcccccaagg ccacactgat tctcaattcc atcggcagcc
901 tcagcaaaact ccaggcccag cctctcacct totcccggtg tggtggcagg ccacagggcc
961 tggctatttc ctctoagcat ggggaagcgc aagtttcgcc tgcctcacc ccaagaaacc
1021 ggaggaaatc agctcataaa gtgttggcag agctagatga cgaggtctcc ccaggccccc
1081 tgtccgtcct gacggagccc aggaagaggc tccgggtgga gaaagctgct ctgatatcca
1141 gtggggaatg accacaggga aaggtccatt gctaaagact gactcaaggo tggagaatgc
1201 ttgcaaatga agtgacagaa aagatggttt cctgccaaag atgttgacgc ctccaagttt
1261 ccagtgcact acaagcccag tcacaaggaa taaagccttg cctccatgga tgccaggata
1321 cctgcagtta gagcaacggc cccagaaaac attcctagaa ggtatggttt tattttgctg
1381 aagccagcat ataccaggca agtgcctctg cacaataccc taacctact catgttcttg
1441 aatggcatcg gctaaagcac cctgaaggac ccagagcaca ttccaaggo atcctcatgc
1501 caccgctagc tagtatcac accagctgag totctgggac tagcttttca gagcttagtc
1561 tttattocca aataaagaat aaactgttcc atccaaaaaa aaaaaaaaaa aSEC ID N°: 34

```

SEC ID N°

39
40
41
42
43

```

-----GGG-GTG-----A Mayoria
1 -----CGG-----G-----As 4507170-SPARC.seq
2 -----GTC-----A 213614-QR-1.seq
1 -----CGGCGATG-----A 21359870-Revlin.seq
1 AACCGCCAGACCGTCCGCGCGGCGTATGCGAGGASCACA 15451924-SPOCK.seq
1 -----CGCGCGCGTATGCGAGGASCACA 13242264-FRP.seq

GAGGC-----CGGCGCTGCC-GGCGA-----ACC Mayoria
10 20
6 GAGCG-----CGGCTCTGCGTCCG-----C 4507170-SPARC.seq
3 GA-----CTTCCTCC-----A 213614-QR-1.seq
9 GAGGC-----CAGCCTGCGCGGAA-----ATC 21359870-Revlin.seq
41 GCGCGCGCGAGCCGCGATCGCGCGTATGCGAGGASCACA 15451924-SPOCK.seq
2 -----TGGCGCTCG-----A-----ACG 13242264-FRP.seq

-CAGGAGCCTG-CACCAAAAACGAGGC-----CTG Mayoria
30 40
27 TCGCTGCGCTG-CAAC-----TAGGG-----TTC 4507170-SPARC.seq
17 CCGGAGCTCA-CAG-AAAAAC-ACG-----CTT 213614-QR-1.seq
32 CAGGAGCTG-CAACAAAACGATG-----CAG 21359870-Revlin.seq
81 CCGCGAGCTCGACCACTCGGACTAGCGGACCGCGCGCG 15451924-SPOCK.seq
14 GAC-TGCTTC-CATC-----CTG 13242264-FRP.seq

CTGCAAAATTC-----CTGCGCGCGAGATG Mayoria
50 60
50 CAGCA-----CC-----ATGAGGCT-----C 4507170-SPARC.seq
42 TCGGAAAGATTC-----ACACTCAATG 213614-QR-1.seq
60 TCGGAAATCTCT-----CTGCTGCCAACCTC 21359870-Revlin.seq
121 CCGCCAGATGCGCGCGATCGCGGATATGCGCGCGCGCG 15451924-SPOCK.seq
30 AC-CAGGCTCT-----CT-----CGAGCTTC 13242264-FRP.seq

GGGA-----TCCG-----GTCTGTCACCTC Mayoria
70 80
67 TGA-----TCTTC-----TTTCTC----- 4507170-SPARC.seq
68 CAA-----TCTTC-----ATATACCAAGG 213614-QR-1.seq
87 CAA-----TCTTC-----GTCTGTCACCTC 21359870-Revlin.seq
161 CCGCGCGCTGCTCTCTCGAGCTCGAGAGCGCGCACCTG 15451924-SPOCK.seq
50 GGA-----CCG-----GAGCAAGATGTC 13242264-FRP.seq

GAAAC-TTTCGAGGCTGCGCGG-CTC--CAGGCGATT-- Mayoria
90 100
82 -----TTTCTCTGCG-----AGGCGCTT----- 4507170-SPARC.seq
92 AAAAT-TTTCGAGGCTGCGCGGATTT-----CATGCGACG 213614-QR-1.seq
108 AGACC-CCGCTAG-TTTCAGAGCGAG-----CAAAATATC- 21359870-Revlin.seq
201 GCGCGCTCGCGCGAGCGCGCGCTCGAACCAAGCGCAAT 15451924-SPOCK.seq
69 GAGC-GCT-----GCGTGGCG-----CGG-----CGCGCT- 13242264-FRP.seq

-----AGTGGCGAGTCTGACTAGAA-CTGCTCTGCTC Mayoria
110 120 130
105 -----GCGAGCGCTCTGAGCAAGAA-----TGC-CTGCG 4507170-SPARC.seq
117 -----AGTGGCTAACCAAGTAGGA-----GCGAGCTCTTC- 213614-QR-1.seq
142 -----AAGT-CAAGTACAGTGAAT-CTGCTCTGCGG 21359870-Revlin.seq
241 TCTAGACAAAGACCAAGTGGCTGAGCAAGCTCTGCGCAAT 15451924-SPOCK.seq
83 -----GCTGAGC-----AT-----CGGCTGCTC 13242264-FRP.seq

```

Figura 13

132 T G A T - G A G A C A - G A A G T G G - - - - - T G G A A C A - - - A C 4507170-SPARC.seq
 146 - - - - - T A - T T G C C A - - - - - T S A A A G G A - - A A 213614-QR-1.seq
 172 C A A A - G A A G C A - G A C C T A A - - - - - C G A G G A A A - - G G 21359870-Havin.seq
 201 G G A C C G G G A C A A G T A C T G G A A C C G C T T T C G A G A C G A T S A T 15451924-SPOCK.seq
 112 C A - - - - - - - - - - - C G G - - - - - C G A G G A G G - - A A 13242264-FRP.seq
 T A T T T A A G A A G T - - - - A T C T C - - T G A G S T A T T T G T C A A A G Mayoria
 160 170 180 190
 159 T G T G G C A G A G G T - - - - G - - A O - - T G A G S T A T T T G T C G G A G 4507170-SPARC.seq
 164 C G T T C A A A C C C T - - - - A C C A G - - T G A G C C A T T A A C T A C C A 213614-QR-1.seq
 199 G A T T T A A A G A G T - - - - T T T T C - - T T G G S T G T T T G T C A A A C 21359870-Havin.seq
 321 T A T T T C A G A A A C T G G A A T O C C A A C A A G C C C T T T G A C C A A G 15451924-SPOCK.seq
 127 C A - - - - - A A G A A G C - - - - A A A T C - - C A A G - - A T C T G C - - - - 13242264-FRP.seq
 T C C T G - - - - - - - - - - T T C C C T G T C T G A - - - - - A A T G Mayoria
 200
 191 - - C T A - - - - - - - - - - A T C C - T G T C C A G - - - - - G 4507170-SPARC.seq
 198 T C C T G - - - - - - - - - - T T T A A G A T C T G A A - - - - - A A A C A 213614-QR-1.seq
 233 T T T T A - - - - - - - - - - T T C C C T G T C T G - - - - - T G 21359870-Havin.seq
 361 C C C T G G A A C C A T C C A A G A C C C T G C C T G A A G S T A A A A T G 15451924-SPOCK.seq
 151 - - - - - - - - - - - - - - - - G C C - - - - - A A T G 13242264-FRP.seq
 T G G - - - - - A G A G T G T A T - T G A A C T T G A A T T - T C - - - - Mayoria
 210 220
 206 T G G - - - - - A A G T A G C A - - - - - G A A T T T G A - - - - - 4507170-SPARC.seq
 221 T G A - - - - - A G A C C G A T - T G C T C C T G A T T T G T C - - - - 213614-QR-1.seq
 251 T G C - - - - - A G A G C G G A T - T C A A C T T C A A T T T T C - - - - 21359870-Havin.seq
 401 C A G C C C T C A C A A A G T G T G T G T G A C C C A G G A C T A C C A G A C C 15451924-SPOCK.seq
 150 T G - 13242264-FRP.seq
 - T T C C G T G G A T C T G G - T C C A G C C C T - - - - - - - - - - Mayoria
 22 23
 229 - - T S A T G C - - T G - - - - - C A - - - - - - - - - - 4507170-SPARC.seq
 248 - T T C T A - G G A T C T G C T T G C A G C A C T C - - - - - - - - - 213614-QR-1.seq
 278 - T G C A G T G A C C C C T G G G T C C A G C C C C C - - - - - - - - 21359870-Havin.seq
 441 G C C C T G T G T G T C A A C C C A A G C A C C T G C T C C C A G G C A A A 15451924-SPOCK.seq
 160 - T T T T G T G A G C T G G - - - - - - - - - - - - - - - 13242264-FRP.seq
 - - - - - - - - - - - T A - C C - - C A A - C A C T G G G - - - - - - - - Mayoria
 235 - - - - - - - - - - - G A - - - - - G C A - - - - - - - - - 4507170-SPARC.seq
 272 - - - - - - - - - - - T A A C C - - G A T C C A T T G - - - - - - - - 213614-QR-1.seq
 303 - - - - - - - - - - - T A C T T - - A A A G A T C T G G A - - - - - - 21359870-Havin.seq
 481 A G A A G G G G A C G T G G C C A G A A A C A C T G G G T T T G A C C T T C 15451924-SPOCK.seq
 174 - 13242264-FRP.seq
 - A A C T - - G T C A A G T G G G G C C T T T T C G C G T - - - - - A A A G T C C Mayoria
 240 250 260
 240 - A A C C - - G - - A A G A G G A G C - - T G C T G C G G - - - - - A A A A T C C 4507170-SPARC.seq
 296 - A A C T - - A C C A A T T T G G G C C C A T G C G A C A G - - - - - A A A A C T C 213614-QR-1.seq
 319 - A A G C - - A T G A A G A C T G C G C C T T T T T T C C T - - - - - A T G T G T C 21359870-Havin.seq
 521 G A A A T T G G T C A A G T G C A A G C C C T G T C C C T G C C A G A G T O A 15451924-SPOCK.seq
 179 - A A - T - G C - G T C A 13242264-FRP.seq

70 250
 270 -CTG---GAGAACG---AT---CAGTG-CGAA-CACGGC 4507170-SFARC.seq
 320 -CAGACAGAGCATAAAT---ATACTCATTCCTGAATGCTAG 213614-QR-1.seq
 353 -TTGGGAACTGCGAGCT---GCAATC-CCGACAAATGCG--- 21359870-Havin.seq
 561 GCGATGCTCTGCGGCTCAGATGGCCACTCCTACACATCCA 15451924-SPOCK.seq
 191 -CGGAGAA- 23242264-FRP.seq
 AG---GATAGAAE-C-ATG-TTG---CAA- Mayoria
 90
 299 AA---GCTG---TG---CGAG--- 4507170-SFARC.seq
 358 AG---GAGAGAAACACAGGGGTTG---TAAA- 213614-QR-1.seq
 385 AA---GATTATTATCTGATCATTG---CAA- 21359870-Havin.seq
 601 AGTGCAAAATTGGAGTTCCATGCTTGTCTACTGGCAAAAG 15451924-SPOCK.seq
 199 GG---GG---GAG- 13242264-FRP.seq
 ---CGAAC-GGTGAGATG---TGTC-C---CTCACA Mayoria
 110 120
 310 ---CTG---GATGACA---ACA 4507170-SFARC.seq
 381 ---CAA-GGTGATGTTG---TGCTG---GCCACA 213614-QR-1.seq
 410 ---CCAACTGCTGAAACGG---TAGCAC---CTGACA 21359870-Havin.seq
 641 CCTCGCCACCCCTCTGTGATGSSGCTGCTCCTGTTCTCCA 15451924-SPOCK.seq
 205 ---CCAAC---GTG- 13242264-FRP.seq
 GCAAC--AA-CCCCAGTGCATGTGCCGAA---CT Mayoria
 330 340
 323 ACACCC---CCCATGTGTGTTGTGCCAGGA---CC 4507170-SFARC.seq
 407 GGAACCATAA---AGCAAGAGGTACCGGTA---CT 213614-QR-1.seq
 439 ACACCTGCAATGCCAGTTTATGCGGCTGAAAG---CT 21359870-Havin.seq
 681 GAGCCTGAGCCGACCAAGCACAAGGCAGAAAGGAGTGCT 15451924-SPOCK.seq
 213 ---GCTCTGC- 13242264-FRP.seq
 GCACAACTGAGAGATG---AACCCAGCCGCCA-CAGAA Mayoria
 350 50 37
 380 CCAACAGCTG---CCAGCGCCCA-TTGG- 4507170-SFARC.seq
 435 GCAACACAGAAAGATG---AGCCTGCTCTC-CAGAA 213614-QR-1.seq
 470 GAAAGAAAGCAAAAGAA- AACAGCAGTATCCA-CAGAA 21359870-Havin.seq
 721 GCACAGAACAGGAGTTGCTGAGACCTGCTGCTGCTGAA 15451924-SPOCK.seq
 220 ---ATTGAGCAATG---CAAACCTCA-CAGAA 13242264-FRP.seq
 GG---CGAGTCTGAGCCTGAGGCTGAGCA---GTC Mayoria
 300 390
 376 ---CGAGTTTGA---GAAGGCTGTGCA---GCA 4507170-SFARC.seq
 470 GA---CAAGCACAGG-GTGATGGTGAAGCA---TCA 213614-QR-1.seq
 506 CA---GCAATGCGAACCATAGGCTGAAAA---ATC 21359870-Havin.seq
 761 GCAATTGGTTTGGAGCTCTCCACGAGGATGCAACAGAGTC 15451924-SPOCK.seq
 245 GC---C---CTGTGTGTG- 13242264-FRP.seq
 ATCAE-----TAATACAGCC-AAGGA----- Mayoria
 100
 399 AT-----GACA----- 4507170-SFARC.seq
 496 AACAA-----AAR-KCAGCCTGAGGA----- 213614-QR-1.seq
 535 ATCAG-----TACTAAAGTGAAGAA----- 21359870-Havin.seq
 801 ATCAAGCCCAACAGCTCGAACACAGCCCAAGGCAAGTTT 15451924-SPOCK.seq
 250 -CAG-----TAAT-----G----- 13242264-FRP.seq

-----GCAAAGAGCT-----TG Mayoria
 410
 404 -----ACAGAGAGCT-----TC 4507170-SPARC.seq
 518 -----GCATTAAAT-----TC 213614-QR-1.seq
 556 -----GGAAAGCCA-----TG 21359870-Revlin.seq
 842 ACACATACCATCCTGCCCATCTGCAAGGACTCCCTGGGCTG 15451924-SPOCK.seq
 256 -----GCCAAGACCT----- 13242264-FRP.seq

GATACACAGCA-A-----ATAA-CCA-GA-TTCCTGCT- Mayoria
 20
 415 G--ACAC-----TTCCTGCC- 4507170-SPARC.seq
 529 CTTACTCTGACACA-----GTAAATCCAGGGTTGGCTCTCT- 213614-QR-1.seq
 567 ASCATTCAGCAGCA-----ACAGGGCAAGAGTTCTAGCC- 21359870-Revlin.seq
 881 GATCTTCAAAOAAATTTGGACATGAACCATGACCTCTGCTT 15451924-SPOCK.seq
 275 ----- 13242264-FRP.seq

-----GAGA-----ACT-GGA-AA-ACG Mayoria
 43
 428 -----A-----CT----- 4507170-SPARC.seq
 562 -----GATA-----ACGAGGA-AA-GCA 213614-QR-1.seq
 600 -----AAGA-----GCTGGCATTGA-AGG 21359870-Revlin.seq
 921 GACCCCTTCAGAGATCAATGCCATCTACCTGGATAAGTACG 15451924-SPOCK.seq
 275 ----- 13242264-FRP.seq

ATCCTTGCACCA-----BTGAGG-TCAGT Mayoria
 440
 431 -TCCTTSCCACA----- 4507170-SPARC.seq
 578 ACTCTGGCAGCA-----ACAGGCAACA- 213614-QR-1.seq
 618 ATCAAGAGGACA-----GTGATGGTCACT 21359870-Revlin.seq
 961 AGCCCTGTATCAAGCCCTCTTTCAACTCGTGTGAOTCCCT 15451924-SPOCK.seq
 275 AGCTCARGCATT-----GTGA-----ACT 13242264-FRP.seq

TAAGTACAGC-----TCTG-----GAGTA-TGG-ACCAO Mayoria
 50 60
 442 -AAGTECA-C-----CCTG-----GAGG-----GC-ACCAA 4507170-SPARC.seq
 600 -----GCACAGC-----TCTG-----AGCA-C-C-ACCA 213614-QR-1.seq
 642 TAAGTGTAA-----TTTG-----GAGTA-TGC-ACCAA 21359870-Revlin.seq
 1001 CAGGGATGCGAAGCTTTCTAACAAATGAGTGTGCTACTGC 15451924-SPOCK.seq
 294 T--CACA-----GAG-A-CGC-----CTGC 13242264-FRP.seq

CTCAAGGGGCACAGGAT-----CCAGGTGCA-ATGA Mayoria
 470 480
 455 -CAAGGGGCACAGGCT-----CCAGGTGCA----- 4507170-SPARC.seq
 522 CCAGGGAGGCACAGGAA-----ACAGGGCAACATGG 213614-QR-1.seq
 659 CTGAAGGTACATTTGGAC-----ATAAAGCAAGATA- 21359870-Revlin.seq
 1041 TTCCAGAAACCTCGGAGGTCTCCCTTGCACAAATGAATGA 15451924-SPOCK.seq
 310 CTCACTSGATCCAAAT-----CCAGGT-----TGA 13242264-FRP.seq

GTAG--TTAATGGGCTTTGCAAGGAGAAA-TCCCA--G-G Mayoria
 490 500 510
 489 GTA-----CATCGGGCCTTGCAGATACA-----TCCC----- 4507170-SPARC.seq
 653 CTGG-----CAATGGGCTCTGAGAGGAGAAAATC-CA--GTG 213614-QR-1.seq
 699 -TGA--T--TGAGCCTCAGGAGAAAAAATCTCTCA--GAG 21359870-Revlin.seq
 1081 ACAGAAATTCAGAACTGAGTAAGGGGAAAGCCCTGTTGG 15451924-SPOCK.seq
 336 TTA-----TGAATGGGCACTGCAGAGGAAA-----G 13242264-FRP.seq

516 ---CGCTTCCTGCGACTC-TGAGCTGACCGAATTG--- 4507170-SEARC.seq
 588 GATGCTCTTSGCTCTGCTTCG-TGAGCGCAACACAT--GGA 213614-QR-1.seq
 731 AACACAGGATTTTTTGGCTCC-TGGTCTTACTTCCCTCACA 21359870-Revin.seq
 1121 GCGCTTCTATACCTCGGTGTAAATGAGGAGCGCTATTACAAA 15451924-SPOCK.seq
 351 AA-----GTCTG-TGAG-----TGCAATCCGCC 13242264-FRP.seq

 GACCC-----CAAGCAATGCG-EGGCTAGC-CAGATG-A Mayoria
 547 ---CCG-----CTGCGCATGCG-EGACTGCG----- 4507170-SEARC.seq
 725 AATACT-----AATAAAAAACA-CAGTTGCGCTAGATGAA 213614-QR-1.seq
 770 GATTC-----TAACTAACCAAG-AAAGTATCACAAAGGA 21359870-Revin.seq
 1161 GCCACACAGTGCACCGGACCAAGGGGCACTGCTGCTBTG 15451924-SPOCK.seq
 382 AGCCG-----CGTGTCTGCG-----TATG----- 13242264-FRP.seq

 -AGGCAAAAT-CGGCAATGAGCTGG-----AGGAAAT Mayoria
 563 ---T-CAAGAACG-----GCTG----- 4507170-SEARC.seq
 757 AACACAAAT-CAAGATCAAGAG-----AGGAGCT 213614-QR-1.seq
 903 GAGGAAAAAC-CAAGAACCACTAG-----AATATTAT 21359870-Revin.seq
 1201 TGACAAAAATATGCAATGAGTGTGCTGGGTCCAGGAAACA 15451924-SPOCK.seq
 401 -AGGCTAAC-CGTGA-TGAGCTGC-----GGCG-----C 13242264-FRP.seq

 CGG-CTCAACAGCTG--AAGAGGAGCAGTGAA----- Mayoria
 580 CCG-CTCAGC--CTG--TATGAGAGCGATGA----- 4507170-SEARC.seq
 787 GCG-G-----AG-G--AAGAGGATGAG-GAA----- 213614-QR-1.seq
 934 CCG-ATCATCAGTGS--AAGAGGAGGAGTAA----- 21359870-Revin.seq
 1241 GGGTCTGTGTGAGCTGTGAAGAGGAGCAGGAAACCTCAGGG 15451924-SPOCK.seq
 427 CCG-ATCATG-----EAGTGGC----- 13242264-FRP.seq

 -----TGGGCGCAAGGAGCTG---ATGACC Mayoria
 506 -----GGACAAACAAECTT---CTGAGT 4507170-SEARC.seq
 909 -----TGCGCTGAGAGCAACTG---ATTAC 213614-QR-1.seq
 953 -----CATAGTCAAGGCTAA---GGGATG 21359870-Revin.seq
 1281 AATTTTGGCAGTGTGCGTCTCGTCTCGTCTGATGAC 15451924-SPOCK.seq
 443 -----TGCAA-----GGC 13242264-FRP.seq

 GAGAAATAGAA--GAGCAGGCTCAAAATGCT----- Mayoria
 525 GAGAA-----GCGAAGCTG----- 4507170-SEARC.seq
 829 AGGCAATAGAA--ACACAGAGCCCG--TGGA----- 213614-QR-1.seq
 895 AAGCAAAACAA--GAGCAGGATCCAAATATTT----- 21359870-Revin.seq
 1321 TAGAATATGAAACGGGAGCTGGGACCAAGGACAAAGAGGG 15451924-SPOCK.seq
 451 GAGATCAT-----TCCAGATGGCT----- 13242264-FRP.seq

 -----CGACTGCTCA--AGGAGAT Mayoria
 540 -----GCGGTGAAGA--AGATCCAT 4507170-SEARC.seq
 957 -----CAAGCCATGCG--AAGAGAT 213614-QR-1.seq
 915 -----CCATGGAGA--AGAGGAG 21359870-Revin.seq
 1361 GAGGCTGAGGCTGACACACCGAGCCCTGACAGAGGATCAT 15451924-SPOCK.seq
 470 ----- 13242264-FRP.seq

560 570 580
 558 GACGAATGAGGAAGCGCTT-----TGCAGCGGACGCGAC 4507170-SFARC.seq
 573 ACAGAAACATGCGCAAAA-----TGAAGAACAGCGCGGC 213614-QR-1.seq
 593 GAGAAAAAGAGCCAGG-----TGAAGTTGGT-ACC 21359870-Havin.seq
 1401 GAGGATGAGGATGACAAAGAGGATGAGGTCGGGTACA 15451924-SFOCK.seq
 470 -----GGTCTCTC-----TAAAGGCGAGT-AAC 13242264-FRP.seq
 GACCCCTGGTGAACAC-AGAGAGA-----TAC Mayoria
 590 700
 688 CACCCCG-TC-----GAGC-----TGC 4507170-SFARC.seq
 905 CTC-TCGTGAATTCCTTGAGAGA-----TTG 213614-QR-1.seq
 962 CACAAATGATAACCAAGAAAGAAA-----GAC 21359870-Havin.seq
 1441 TATGATATGTCGCCACRAGAAAGACGACACAGATTTTGAC 15451924-SFOCK.seq
 490 TACAGTG-----AGA-----TCC 13242264-FRP.seq
 TAGATTG-----CCAGTACTTATA-CTAA Mayoria
 710
 704 TG-----GC-----CGGGGACTT-----CGA 4507170-SFARC.seq
 930 CAGTCTG-----C-AGTACGTATAACCA 213614-QR-1.seq
 988 AGAATTG-----CCAGGGAAGCATG-CTAA 21359870-Havin.seq
 1481 AAAATTTGCAAGCTCACTTCTATATCTGCTGATTTGATCTAA 15451924-SFOCK.seq
 503 TAGA-----CAAGTACTT-----TAA 13242264-FRP.seq
 GA-----GCGACCATCA----- Mayoria
 720 730
 720 GA-----AGAAATATAT----- 4507170-SFARC.seq
 954 GA-----GACACGGGTGA----- 213614-QR-1.seq
 1012 CA-----GCAAGCAGGA----- 21359870-Havin.seq
 1521 GATCCAAAGGCAAGAGGTCTCTCTCTCATTTGTTGCTCTC 15451924-SFOCK.seq
 519 GA-----GC----- 13242264-FRP.seq
 -----GGAAATACAACT-----TCC Mayoria
 740
 732 -----GATGTACATCT-----TC 4507170-SFARC.seq
 966 -----GAAATTCAGCA-----TGC 213614-QR-1.seq
 1024 -----GGAGACAAATA-----TGC 21359870-Havin.seq
 1551 TATACCGACCTAAGGTTTGGAAAGCAACTGCTTGTTC 15451924-SFOCK.seq
 523 -----TTCGATG-----ATATTT-----TGAAGAG----- Mayoria
 750
 745 -----CGG-----TACAG-----TGGCA----- 4507170-SFARC.seq
 990 AG-----CAGG-----AAAGT-----AGCAAAA----- 213614-QR-1.seq
 1039 AA-----TCTGATG-----ATATTT-----TGGAGAG----- 21359870-Havin.seq
 1601 AGAGGATTTGTGATTTTTCATATGTTTGAATGGGAGAAAG 15451924-SFOCK.seq
 529 -----TTTGATA-----ATGGT-----GA----- 13242264-FRP.seq
 -----TTCGGTC Mayoria
 760
 760 -----TTCGGTC 4507170-SFARC.seq
 1000 -----CTCTA-C 213614-QR-1.seq
 1061 -----TCTGATC 21359870-Havin.seq
 1562 GTGTGTGTGTTTTTTTTTTGTGTGTGTATTTTGTGAT 15451924-SFOCK.seq
 537 -----CTC-----TC 13242264-FRP.seq

AGGCTAGTCA-----AGCAGGCT----- Mayoria
 770 780
 767 AGCTGGACCA-----GCACCG----- 4507170-SPARC.seq
 1006 AAGGAAGGAA-----AAGTCCCT----- 213614-QR-1.seq
 1068 AACCAACTCA-----AGTAAGCA----- 21359870-Havin.seq
 1581 AAGCAAGTCATTGGCTTAATTAAGACCTCCTCCCTTTCTG 15451924-SPOCK.seq
 542 A-----CCT----- 13242264-FRP.seq
 -----AGACTTC-----AGCGGATGATTTTCATGAGGTTTC----- Mayoria
 790 800
 783 -----CAFT-----GAGGGGTACCTCTC-CGA-----C- 4507170-SPARC.seq
 1024 -C--TCTC-----AAGGAAATTCATATGAGGATC- 213614-QR-1.seq
 1085 --AGATGCA-----GGAGGATGAATTTGATCAGGCTA- 21359870-Havin.seq
 1721 TGAATTTTTCACAAACATGCTGATTTACGTGGAATTCCT 15451924-SPOCK.seq
 546 --GGACTCC-----AGCGAAT-----TCTGAAATTC- 13242264-FRP.seq
 -----ACCGAGGAG-----AAAAG- Mayoria
 810
 805 -ACCGAGG----- 4507170-SPARC.seq
 1052 -A--AGGTC-----AAAAA- 213614-QR-1.seq
 1116 -ACCAAGAAC-----AAGAA- 21359870-Havin.seq
 1761 GACAGTGCAGGGAGCCCGACCCCTCTAAATGTCAAAGAC 15451924-SPOCK.seq
 571 -GTGGAGGAG-----AATGA- 13242264-FRP.seq
 -----GATAACTC-CACTG-CAGTAA--ACAGCAT-GT Mayoria
 820
 812 -----TGGCTC-CACTG-C----- 4507170-SPARC.seq
 1064 -----GACAAAGT-GAAGA-AGGTAA--AGAGCATTTT 213614-QR-1.seq
 1130 -----GATAACTC-CAATG-CAG-AA--ATGGAGAGAG 21359870-Havin.seq
 1801 CTTTATGATTACCGACACTGGTGGTTATTAAGGATAGG 15451924-SPOCK.seq
 586 -----AACAGCG-----GTCA--ACA----- 13242264-FRP.seq
 -----AAGTCTTCTCTCAT-----CG- Mayoria
 830
 824 -----GTGCTCCTCTCAT-----C- 4507170-SPARC.seq
 1098 C-----AAGTA-GTCAATCAG-----CG- 213614-QR-1.seq
 1159 A-----AATCTCTCGAAGCGT-----CAA 21359870-Havin.seq
 1841 TCCAGCCCTTACAGTGTCTAAGTGGCTCTCTGTGTCCTG 15451924-SPOCK.seq
 599 -----TCAAG-----CGC 13242264-FRP.seq
 TAAACCC-----AT--AAGAAAC-AAAAG----- Mayoria
 840
 838 -----CCC-----AT----- 4507170-SPARC.seq
 1110 CAXAUCAC-----AGAGCAGTGACGAAAG----- 213614-QR-1.seq
 1177 TAAAGCAC-----ATCCAGGAACTGAATG----- 21359870-Havin.seq
 1981 TAGATGTTGTGAAGGAAGAAACAAATAACACCA 15451924-SPOCK.seq
 606 TTACCGG-----A----- 13242264-FRP.seq
 -----GCAGG----- Mayoria
 843
 1134 -----GCAGG----- 4507170-SPARC.seq
 1201 -----GCAGG----- 213614-QR-1.seq
 1921 GTGTAGTTTTTCCCTTGGCCCGCTACTGCGGCTGATTA 21359870-Havin.seq
 614 ----- 15451924-SPOCK.seq
 13242264-FRP.seq

-----GAGAG----- Mayoria
 843 -----GAGAG----- 4507170-SPARC.seq
 1139 -----ATAGAGAG----- 213614-QR-1.seq
 1206 -----GTCAAGAG----- 21359870-Havin.seq
 1951 TTATTAAAAATTTAGTTTCTTTCACATCATTATATCTGGCT 15451924-SPOCK.seq
 614 -----ATCAGGAG----- 13242264-ERP.seq

-----GACAACA-----CTGCTCAAGA----- Mayoria
 850
 847 -----CATTGCA-----CCACCG----- 4507170-SPARC.seq
 1147 -----GGCAGCA-----ATGCTGAGGA----- 213614-QR-1.seq
 1214 -----GGTAAA-----CTGGCTAGA----- 21359870-Havin.seq
 2001 TCCATAAAGAACAGCCCTTAATTCAGTCAAGACTGCCCTTT 15451924-SPOCK.seq
 622 -----AACACA----- 13242264-ERP.seq

-----GGA-----TATTAA-AAT----- Mayoria
 860
 1164 -----GGA-----TGATAATGAT----- 4507170-SPARC.seq
 1231 -----AGC-----TATCAGCAGC----- 213614-QR-1.seq
 2041 GCGAATTCATTTTATTAAAAATTTGGTCTCTGGTACTTTC 21359870-Havin.seq
 629 ----- 15451924-SPOCK.seq
 13242264-ERP.seq

-----CATAAAGAG-----ATGGA-AGGAAGAGCTG Mayoria
 860
 1177 -----AGTGGGTGATG-----ATGGTGAAGGAGATCT 4507170-SPARC.seq
 1214 -----CACAAAGAG-----ACAGAAAGAAAGACTG 213614-QR-1.seq
 2001 CTGTACATGCAATTAATATGCAATGCAATGACAGGAAGAGCTG 21359870-Havin.seq
 629 -----AACTG 15451924-SPOCK.seq
 13242264-ERP.seq

-----CGATGCTCTCT----- Mayoria
 870
 864 TTT-----CGAGACCT----- 4507170-SPARC.seq
 1203 TGG-----CAATGCTCTG-----AGCG 213614-QR-1.seq
 1269 TTT-----CTGAGGCTCT-----GCTG 21359870-Havin.seq
 2121 TATGTGCTGCTCTTGCACACACACCCATACCTCTCAGAA 15451924-SPOCK.seq
 634 CTC-----AGAGGCTCTCT-----G----- 13242264-ERP.seq

AAAGTGG-TCTATTGA-----CGATGATAA-T----- Mayoria
 880 890
 875 -----CTGA-CCTG-----CAATGACAGT----- 4507170-SPARC.seq
 1220 AAGCAG-CTTAQ-GA-----GCAAGAGGA-A----- 213614-QR-1.seq
 1286 ATGGA-CCTACTGA-----TGATGCTAA-T----- 21359870-Havin.seq
 2161 AAGTGGCTATATTAAAAACTCGAAAAACATGATAAATTT 15451924-SPOCK.seq
 648 -----TGTGA----- 13242264-ERP.seq

-----AGCTTGGCCAG-----AATGATTGGGCTGG----- Mayoria
 900 910 920
 896 -----AATTCGGCTG-----GATGAGTGGGCTGG----- 4507170-SPARC.seq
 1243 -----AGAAAGCAGAGC-----AATGACCAAGGACAG----- 213614-QR-1.seq
 1310 -----ACGACGCGCCAG-----AATGATGGAGCTGA----- 21359870-Havin.seq
 2201 CTGAGCTTGTCCAGACCTGGAGCAAAATTTCTCGAATAAG 15451924-SPOCK.seq
 654 -----TGGCC-----TCAATGAACTG----- 13242264-ERP.seq

921 507170-SPARC.seq
 1269 213614-QR-1.seq
 1236 21359870-Havin.seq
 2241 15451924-SPOCK.seq
 670 13242264-FRP.seq
 923 507170-SPARC.seq
 1282 213614-QR-1.seq
 1349 21359870-Havin.seq
 8281 15451924-SPOCK.seq
 671 13242264-FRP.seq
 948 507170-SPARC.seq
 1302 213614-QR-1.seq
 1266 21359870-Havin.seq
 2321 15451924-SPOCK.seq
 690 13242264-FRP.seq
 987 507170-SPARC.seq
 1333 213614-QR-1.seq
 2397 21359870-Havin.seq
 2351 15451924-SPOCK.seq
 696 13242264-FRP.seq
 997 507170-SPARC.seq
 1163 213614-QR-1.seq
 1424 21359870-Havin.seq
 2401 15451924-SPOCK.seq
 721 13242264-FRP.seq
 1021 507170-SPARC.seq
 1399 213614-QR-1.seq
 1453 21359870-Havin.seq
 2441 15451924-SPOCK.seq
 738 13242264-FRP.seq
 1419 213614-QR-1.seq
 1481 21359870-Havin.seq
 2481 15451924-SPOCK.seq
 746 13242264-FRP.seq

AAGAGGTGCHXACA-AATTTGCTGTTCT-----GGTAGCA Mayoria
 1040 1050 1060
 1036 --AAATSTTTGGATG-GTTTSTTGTGTT-----GG-----CC 4507170-SPARC.seq
 1422 --ACAACTGCAACA-AGTAGCTCTGTT-----GAGAGCA 213614-QR-1.seq
 1485 --AAAAAGTACATGAA-AATGAAATATA-----GGTAGCA 21359870-Revlin.seq
 2521 TAGGCATGCAACACATATCTGCTGTCTCTCAGAGTACCA 15451924-SPOCK.seq
 748 --ACAAAGTGG-----CCCT-----GG-AGGA 13242264-FRP.seq
 TGG-AATGCTACAGAGTACGAAGAT-----GGAGTT Mayoria
 1070 1080
 1064 TGG-AG-----ACAGGTGCTAAAT-----AGATTT 4507170-SPARC.seq
 1454 TGA-ATTCAACAGAGCATCAAGAT-----GAGGTT 213614-QR-1.seq
 1518 CTG-AGCCTGGAGAGCAACCAAG-----GGCC 21359870-Revlin.seq
 2551 TGGTGTCTCTTAAAGAAAGAAATTTCTAATCTCTCAACTC 15451924-SPOCK.seq
 768 CGA-AACCTA-----TGC--AGAT-----GGAGCT 13242264-FRP.seq
 AAGA-----CAACAGAXAAATCATGA Mayoria
 1090 1100 1110
 1090 AAGT-----GAATACCTTAACGGTGC 4507170-SPARC.seq
 1483 AAGA-----CCACAGGAGGTTCAATAT 213614-QR-1.seq
 1544 AAGA-----AAGCAGAGAACTCATCA 21359870-Revlin.seq
 2501 AATGTTTTGCTTTTACTCCCTTTCTACCGACAAATCATGA 15451924-SPOCK.seq
 790 GAA-----C----- 13242264-FRP.seq
 TATGAG-----GTGGA----- Mayoria
 1120
 1111 TA-AAA-----ATGAA----- 4507170-SPARC.seq
 1504 CATGAG-----GA--A----- 213614-QR-1.seq
 1565 AATGAG-----GAGGA----- 21359870-Revlin.seq
 2641 TAGGGCCACAAAGCTGTACAGGATTTTTTTTTTTTAAACCAC 15451924-SPOCK.seq
 795 --CGAG-----GTGGA----- 13242264-FRP.seq
 --AATGCCAAGTGAGGGC-----A-GGTGA-GGC Mayoria
 1130 1140
 1121 --AATTTTAACCCAAAGAC-----A--TGA----- 4507170-SPARC.seq
 1513 --AGTG-CAAGGAATGAGC-----ACTGGGAAGGC 213614-QR-1.seq
 1576 --AACGTCAAGTGAGGC-----AACATGAGGGT 21359870-Revlin.seq
 2681 TCAATGCCAATGGAGGCCCTACAAAGACATCGTAATAAG 15451924-SPOCK.seq
 804 ----- 13242264-FRP.seq
 CAT-----CCCTGGATTTTTC-AGXAACTTC-- Mayoria
 1150
 1142 CAT-----TCTTAGC-TGTAACCTTA-- 4507170-SPARC.seq
 1539 TC-----TCCCGCATCTTTGT-ACAAACTTC-- 213614-QR-1.seq
 1603 CATG-----CTGTGSAITCTTGC-ATGAGCTTC-- 21359870-Revlin.seq
 2721 AGATGGAAGCAAAACCTGGCTTTTAAAGAGCAAAATCTGT 15451924-SPOCK.seq
 804 ----- 13242264-FRP.seq
 -----CACT-----GCAAAA-----GAGGA Mayoria
 1160
 1160 -----ACT-----ATTAAAG-----G----- 4507170-SPARC.seq
 1564 -----CACT-----GCAAAA-----GAGGA 213614-QR-1.seq
 1631 -----CGGT-----GTAAA-----GAGGC 21359870-Revlin.seq
 2761 CCCCCCTCACTCCCGCAAGTGCACAAAGATACTAATGAAGA 15451924-SPOCK.seq
 804 -----CT-----GCA----- 13242264-FRP.seq

170 1180 1190 Mayoria
 1170 CCTATTCCACACGCA TTAACAGG 4507170-SFARC.seq
 1579 AAGGTCTGCGACGCA CAAGCAAG 213614-QR-1.seq
 1646 CAGATCTGTAAGGCAGA CCAACAGG 21359870-Havin.seq
 2801 AAGTTCTTCACCATAGG GTTGTCTTAACTAAACCTATG 15451924-SPOCK.seq
 B10 13242264-FRP.seq
 AATCTCGCTGTA TTTCTCAGCATC CTGC Mayoria
 1200
 1192 TCC CA TTTCTG TC TGC 4507170-SFARC.seq
 1606 AAACCCAGCTGTA TTTGCCAAGATC CTGC 213614-QR-1.seq
 1673 AAACCTCAGCTGTG TCTGCCAGGATC CAGT 21359870-Havin.seq
 2841 GAGTCTAGTTCCAAATT TGGGTAGGGTCACTATCTCTACAT 15451924-SPOCK.seq
 810 TCGCTGTG TCGGT TC CTGT 13242264-FRP.seq
 GATTTCCTCTGACCAAC GTCTTGAT CG Mayoria
 1210 1220 1230
 1209 CATTTGTAGCTTTGCCCAT TGTCTTAT TG 4507170-SFARC.seq
 1635 TGCTTCCTCTGCCACCAAG ATTATAAG CG 213614-QR-1.seq
 1702 SACTTGTCTCCAAACAAAC CCTTGTAT CA 21359870-Havin.seq
 2861 TCTTAGGATTTCTCTCGCTATCAAGCTGCCCCAGATACA 15451924-SPOCK.seq
 829 CG AC AC TG 13242264-FRP.seq
 AGTTTGTGGGACTGACAAATCGGA CCT Mayoria
 1240 1250
 1238 GCACTGG GTGGCA CGGA TCT 4507170-SFARC.seq
 1655 GGTTTGTGGGACTGATATAAGG CT 213614-QR-1.seq
 1732 AGTTTGTGGGACTGACATCAGA CCT 21359870-Havin.seq
 2921 AGTACCAAAAGTAGTCTCTGTGAAGTTCCCCCTTCTCTTC 15451924-SPOCK.seq
 837 GGTCTT GAC AGCGAT GA CCT 13242264-FRP.seq
 GTGTTGTTCTCTGG CAACTTTCGGGCACTAAAT GT Mayoria
 1260 1270 1280
 1260 GTGGGTTCTCTGC G TTAAC AC 4507170-SFARC.seq
 1691 AGATGCTACCTGCG CAGCTCTTTGGCACCAAT GT 213614-QR-1.seq
 1758 AGCTAGTTCTCTGT CATCTATTCTGCTACTATAT GT 21359870-Havin.seq
 2961 AGTACCGCTCTATAGCTACTGTCCGCCACTGATTTTCTAT 15451924-SPOCK.seq
 857 GTGATGG 13242264-FRP.seq
 CXATTGAG GCGATCAG AAGGGG Mayoria
 1261 ACATTT GAG 4507170-SFARC.seq
 1726 CAACCTGAA GGGACAAA AATGGGA 213614-QR-1.seq
 1793 AACTGGA GGGACAAA AAAAGGG 21359870-Havin.seq
 3001 CTATCAGGCTCTCTATATCAGATTCAGCCACCCAAAGGAG 15451924-SPOCK.seq
 864 AA AGATCAG AAGGGG 13242264-FRP.seq
 CXTC Mayoria
 1290
 1290 CTTC 4507170-SFARC.seq
 1750 CAGC 213614-QR-1.seq
 1817 CATC 21359870-Havin.seq
 3041 CCTCTCTGGCCCACTATCTATCTCTGCTTCCGCCATG 15451924-SPOCK.seq
 861 13242264-FRP.seq

AACTTCAGCTGCTTATATATGCTGCC--TGCAAACTCA--- Mayoria
 1294 AACTTCAGCTGCTTATATATGCTGCC--TGCAAACTCA--- 4507170-SPARC.seq
 1756 AGCTGCAGCTGCTTATATATGCTGCC--TGCAAACTCA--- 213614-OR-1.seq
 1821 AACTCCAGCTGCTTATATATGCTGCC--TGCAAACTCA--- 21359870-Hevin.seq
 3091 AACTTCAGCTGCTTATATATGCTGCC--TGCAAACTCA--- 15451924-SPOCK.seq
 891 -----TCCAGACCCA--- 13242264-FRP.seq

 -----TACTTA-CCTGTACA-GATTTTGG-----AGTG Mayoria
 1327 -----TACTTA-CCTGTACA-GATTTTGG-----AGTG 4507170-SPARC.seq
 1790 -----TACCCC-CTGTACT-GATTACGA-----AGTG 213614-OR-1.seq
 1857 -----TTCCTA-CTGTACG-GACTTTGA-----AGTG 21359870-Hevin.seq
 3121 AAGAAATATGTTTACGCCAAACATGATTATATGTCAGCT 15451924-SPOCK.seq
 891 -----CAC-----AAGA-----AG----- 13242264-FRP.seq

 AATGAGTTTCTT-TTCGGGT-----CAGAGACTG Mayoria
 1350 AATGAGTTTCTT-TTCGGGT-----CAGAGACTG 4507170-SPARC.seq
 1816 AATGAGTTTCTT-TTCGGGT-----CAGAGACTG 213614-OR-1.seq
 1833 AATGAGTTTCTT-TTCGGGT-----CAGAGACTG 21359870-Hevin.seq
 3161 AATGAGTTTCTT-TTCGGGT-----CAGAGACTG 15451924-SPOCK.seq
 901 -----CAGA-TG----- 13242264-FRP.seq

 GGTCA-----AGGATATCCC-----CAGGCGGCAATTA Mayoria
 1376 GGTCA-----AGGATATCCC-----CAGGCGGCAATTA 4507170-SPARC.seq
 1845 GGTCA-----AGGATATCCC-----CAGGCGGCAATTA 213614-OR-1.seq
 1912 GGTCA-----AGGATATCCC-----CAGGCGGCAATTA 21359870-Hevin.seq
 3201 ACTCACCTTAAATTAGCATGTTCCAAATCCAGTGGGCATTG 15451924-SPOCK.seq
 907 AC-----CAGATATGCC-----CAGC----- 13242264-FRP.seq

 --TGAA--GGTAACTCCGGAC--ATA-CCTGTTT-- Mayoria
 1400 --TGAA--GGTAACTCCGGAC--ATA-CCTGTTT-- 4507170-SPARC.seq
 1872 --TGAA--GGTAACTCCGGAC--ATA-CCTGTTT-- 213614-OR-1.seq
 1929 --TGAA--GGTAACTCCGGAC--ATA-CCTGTTT-- 21359870-Hevin.seq
 3241 CCTGAATACAGTAGCATCATACCTATAGTTCCTCTTAGAT 15451924-SPOCK.seq
 923 -----AACTCAGAA--GCACCAAG----- 13242264-FRP.seq

 -----CTAAXTGAA-AGCAGAG----- Mayoria
 1416 -----CTAAXTGAA-AGCAGAG----- 4507170-SPARC.seq
 1900 -----CTAAXTGAA-AGCAGAG----- 213614-OR-1.seq
 1967 -----CTAAXTGAA-AGCAGAG----- 21359870-Hevin.seq
 3291 AAGAAATGAACTACTGATATAGCAAGTCTCTTGGCTTC 15451924-SPOCK.seq
 941 -----GAC-AGCAGAA----- 13242264-FRP.seq

 --AAATAAGGCAAGA-----ATATTTACTT-----GG Mayoria
 1431 --AAATAAGGCAAGA-----ATATTTACTT-----GG 4507170-SPARC.seq
 1917 --AAATAAGGCAAGA-----ATATTTACTT-----GG 213614-OR-1.seq
 1984 --AAATAAGGCAAGA-----ATATTTACTT-----GG 21359870-Hevin.seq
 3321 GTAAATAGGCTGAGGTTTGTACTTACTTTCCCAATAGG 15451924-SPOCK.seq
 952 --AA--GACCAAGA-----A----- 13242264-FRP.seq

---GAGAG---TATCTCTTCTGGGG---Mayoria
 1454 --GACTAG---AGGCACAGTGGTGGGA-- 4507170-SPARC.seq
 1943 ATGAGAAAG---CGCTCTCTCTCTGGTA-- 213614-QR-1.seq
 2010 ATGAAAAG---AGCCTCTTCTGGGG--- 21359870-Havin.seq
 3361 AAGACAGACCATAAGCAAACTCTGTTTGGGATCTCAACT 15451924-SPOCK.seq
 963 ---GCTG--- 13242264-FRP.seq
 ---AACACCCCT---TTGATCTT---CTCTTGAAGA Mayoria
 1475 --GAGATCCCT---GCGAGATC---GACC---A 4507170-SPARC.seq
 1967 --AGCAACCAG---TTGAGCTT---CTCTCTGCAAG 213614-QR-1.seq
 2034 --ACCATCCCA---TTGATCTT---CTCTTAAGGG 21359870-Havin.seq
 3401 CCAATCACCTTCTGCTTCAATATTTTCTCTCTGAACTA 15451924-SPOCK.seq
 967 --AACACG---AAGA 13242264-FRP.seq
 ACTTTAAGAAATTTAC---CACATG---TATCTGTATC- Mayoria
 1498 AGCAGAACGTGGTITCC---CG---AGGCTGTAAAG- 4507170-SPARC.seq
 1994 ACTTTGAGAAAACTAC---CACATG---TATCTCTATC- 213614-QR-1.seq
 2061 ACTTTAAGAAACTAC---CACATG---TATCTGTATC- 21359870-Havin.seq
 3441 ARACTGAGATAATTTAGAAAACAGGGCTTAATTGCAATA 15451924-SPOCK.seq
 978 GATCTAAGAAAGAG---GCACG---TA---GACC- 13242264-FRP.seq
 ---CTGTG---CACTGGAAATT---TTGTGAG Mayoria
 1528 ---TGAG---AGAAAG---TTCTGGG 4507170-SPARC.seq
 2027 ---CTGTG---CACTGGCAATT---TTATCAAG 213614-QR-1.seq
 2094 ---CTGTG---CACTGGCAAGT---TAGTGA 21359870-Havin.seq
 3401 AAATTAAGTATGAGTATATTAATAAAATCACGACATTTGATAA 15451924-SPOCK.seq
 1003 ---TCAT---CTGCA--- 13242264-FRP.seq
 ---CTTGGCCAAACACCTATAGATAGA-CATTCTCAC-AT- Mayoria
 1546 --GCTGTCTTATG---AAAATATAGA-CATTCTCAC-AT- 4507170-SPARC.seq
 2050 --CTTGACCAAGCACCCGCTTACAGATCACTGACTU-AT- 213614-QR-1.seq
 2117 --CTTGACCAACACCCCTATGGAATAGTCTTGACAC-AT- 21359870-Havin.seq
 3521 ATCTCACTTTTACATCATCAAGAAACCATTTGTTACTATC 15451924-SPOCK.seq
 1013 ---CCGAGCACCT---CTCTTCA---GC- 13242264-FRP.seq
 ---TCTGAGCTCAG---TCCA---TTGCG Mayoria
 1578 ---AAGCCAG---TTCA---TGACG 4507170-SPARC.seq
 2086 ---TCAGAGCTCGC---TCT---TTGAG 213614-QR-1.seq
 2153 ---TCTGAACTTCT---TCT---CTCG 21359870-Havin.seq
 3561 TCTTCTGAGCTTGGAAAAATGTACAGAGACAAATTA 15451924-SPOCK.seq
 1033 ---GCTAAGCCCA---TATA---C--- 13242264-FRP.seq
 AGCG---TCTCTGGTA---CCTAT-GGAACA Mayoria
 1595 ATTT---CTCTCTTA---CTTT---CA 4507170-SPARC.seq
 2106 AGCG---TCTCTGGTA---CCAT-GGAACA 213614-QR-1.seq
 2173 AGCA---TCTCTGGTA---CCAT-GGAACA 21359870-Havin.seq
 3601 ATGGAATAATGATTTGACTTGAAGAAAGCTTGTAGAAC 15451924-SPOCK.seq
 1049 AGCG---TCTCTGGTA---ATCA 13242264-FRP.seq

TTTTATCTTUCGTTTCTTTTUGGAGT-----GTGA Mayoria
 1615 G--TGCAT-----GTTTC--TTTTCACATTC-----AGGC 4507170-SPARC.seq
 2130 C--TGCATAACCCGTTTC--TTCCAGGAGT-----GTGA 213614-QR-1.seq
 2187 C--TGCATAACCCGTTTC--TTTGAAGGAGT-----GTGA 21359870-Hevin.seq
 3641 GGGTGAACCACTGATTTTAAATTTGCTAATTAATCTTATGA 15451924-SPOCK.seq
 1066 C--CGAATCACCACATATT--TGCT-----T-----GT-A 13242264-FRP.seq

 GG-----GCAACTAAGATAA-ACTTA----- Mayoria
 164 1630
 1639 TG-----TTGGTT-----CAA-ACTTT----- 4507170-SPARC.seq
 2160 TS-----GAGACCAAGACAA-ACTTA----- 213614-QR-1.seq
 2227 CC-----CCAACAAGGATAA-GCACA----- 21359870-Hevin.seq
 3651 CAAGTATCAAATTAAGATGACACTTAAAGATGCTTAGCAT 15451924-SPOCK.seq
 1090 CG-----GCAGCA-----AA-TCCTTA----- 13242264-FRP.seq

 TCACCT--TGA--GAGAGTGGTGC-----CACTGCTT Mayoria
 164 167
 1655 TGG-----GA--GCACGGACTGT-----CACTTCTTC 4507170-SPARC.seq
 2180 TCACCT--TGA--AGGAGTGGTGC-----CACTGCTT 213614-QR-1.seq
 2247 TCACCC--TGA--AGGAGTGGGCG-----CACTGCTT 21359870-Hevin.seq
 3721 TAACTTAATGATGGAGAGAGTGTCTCAACAGACAGTTCC 15451924-SPOCK.seq
 1105 TC-----TGT-----T--TGTTT 13242264-FRP.seq

 -----TGGCAT-----TACG Mayoria
 1680
 1679 -----TGGGA-----AGTG 4507170-SPARC.seq
 2202 -----TGGAT-----TACG 213614-QR-1.seq
 2275 -----TGGAA-----TAAA 21359870-Hevin.seq
 3761 AGTAAGGTRATGAGATGCCATTTTCGGAGACATCTTAAGA 15451924-SPOCK.seq
 1315 -----TGCAT-----AAGG 13242264-FRP.seq

 AAG-----AGGACAT-AAATGAAAATCTCTT Mayoria
 169 700 1710
 1688 GTC-----AGCGCAT-CC-TGAGGGCTTCT 4507170-SPARC.seq
 2219 AAG-----AAGACAT-AAATGAAAATCTCTT 213614-QR-1.seq
 2285 AAG-----AGGACAT-AGATGAAAATCTCTT 21359870-Hevin.seq
 3801 AGATATTTTGAATTCATTAAAGACATTAAATAAAGGCTTC 15451924-SPOCK.seq
 1127 AAG-----TGAGGGTGGCTG 13242264-FRP.seq

 GTT--TTGAG--CCCAGATTGTGGAGAACACA----- Mayoria
 1720 1730
 1712 GCT--CC--TCTCTCTTTTGGAGAAC----- 4507170-SPARC.seq
 2244 GTT--CTGAG--CCAC-CTGAGCAGAAATCC----- 213614-QR-1.seq
 2311 GTT--TTGAA--CCAGGATTTTAAAGAACTCA----- 21359870-Hevin.seq
 3841 CTGAGATTGGAAACCCGAAATCGATGGAGTACATTAATA 15451924-SPOCK.seq
 1142 GCT--AG--CCAGGGCAG-GCAGGCCACA----- 13242264-FRP.seq

 --ACTTT-CAG--CACTCT--C-TG----- Mayoria
 1740
 1734 -----CAG--GCTCT----- 4507170-SPARC.seq
 2270 --CCATGCAG--GCTTACAGCTTG----- 213614-QR-1.seq
 2339 --ACTTTGAG--CATCTCTCTCTG----- 21359870-Hevin.seq
 3861 ATACTTTTCATGCTTCACTTGGACATGACAGCAATTCGATT 15451924-SPOCK.seq
 1166 --ACTTT-----CAC----- 13242264-FRP.seq

TTTTAAAGGCTTCAATGTAATTTAAGGC Mayoria
 1750 1760
 1743 TCTCAGGGGCG - - - - - TCTAGGGGAC - TCC CAGGC 4507170-SPARC.seq
 2290 TC - AAACA - - - - - GCNATSCCCATTATGAC 213614-QR-1.seq
 2350 TTCTAACCAAC - - - - - TTCAGAAATATATGCCAGC 21359870-Havin.seq
 3921 TTTTTTAAAGATCTTTATATACTTTTCRTGAGTACTAAAGAT 15451924-SPOCK.seq
 1171 TTCTAGGAAT - - - - - GC - - - - - 13242264-FRE.seq
 TGTAAT-AGCTA - - - - - GTATGTTTATT Mayoria
 1770 1780 1790
 1778 TGTFT-TGAGCCA - - - - - GSAAGGCCAAA 4507170-SPARC.seq
 2316 TGCAATTAACAG - - - - - CTCTGTTAATT 213614-QR-1.seq
 2368 TGTQAT - ACTT - - - - - GTAGATTTATA 21359870-Havin.seq
 3961 TGTAAAGAGITACCTTATCCTTAGAAAGTAGAAATGTTTGCT 13451924-SPOCK.seq
 1186 - - - - - 13242264-FRE.seq
 TTC - - - - - AGCAATAAGCTTGGCTAGAGTAC-TTGA Mayoria
 1800 1810
 1792 ATC - - - - - AAGAGTGAGATG - TAGAAGGT-TGTAA 4507170-SPARC.seq
 2339 TTC - - - - - AGGARTAACTTGGCATAGCATTC-TTCCA 213614-QR-1.seq
 2409 TTT - - - - - AGCAAAATGTTAGCATGTATGAC-AAAGC 21359870-Havin.seq
 4001 TTTCTCTCTGTTGGAATAGTCTTCCAAAGAGTCCACTTGT 15451924-SPOCK.seq
 1186 TTT - - - - - AAGA - GACACTAAAGGGCAC-CTTGG 13242264-FRE.seq
 AGCAG - - - - - AGAAGTGGCTTGT--ATAACC--TATGT Mayoria
 1820 1830 1840
 1820 ATTAG - - - - - AAAAGTGG - - - - - AGTT--GGTGA 4507170-SPARC.seq
 2370 GGCACT - - - - - AACAACTGGCTCTTG-GATACCA--CAACT 213614-QR-1.seq
 2446 AATGA - - - - - GGTAAATTGCTTGAC-AACAACC--TATCO 21359870-Havin.seq
 4041 ACCAGTGAACAAGAACTCACTTGTATAGTGAACCAAGTACAC 15451924-SPOCK.seq
 1213 GGCAG - - - - - GAG - - - - - GC - - - - - GAGT 13242264-FRE.seq
 AT - - - - - CGGGTGT - - - - - Mayoria
 1850
 1843 AT - - - - - GGGTGT - - - - - 4507170-SPARC.seq
 2402 AC - - - - - CTAATGT - - - - - 213614-QR-1.seq
 2472 AC - - - - - CAGGAT - - - - - 21359870-Havin.seq
 4091 ATAAAAATACCGATTATAAAAATATTTGTAATAAACCACTTC 15451924-SPOCK.seq
 1227 AT - - - - - QGGATTG - - - - - 13242264-FRE.seq
 - - - - - TAGCATTAA - - - - - CTCAGTAAC - Mayoria
 1860
 2452 - - - - - T - GTTT - - - - - CCTCA - 4507170-SPARC.seq
 2412 - - - - - TA - GAA - - - - - ATTGATTAC - 213614-QR-1.seq
 2482 - - - - - TAAcATTAR - - - - - CTTGGAAC - 21359870-Havin.seq
 4121 GTCATTGATACCTGCTATTAAAGCTGAATCCTCACTAATTC 15451924-SPOCK.seq
 1237 - - - - - CAG - - - - - AGAGC - 13242264-FRE.seq
 - - - - - AGAAGT - - - - - G-GCA-TTTGA - - - - - Mayoria
 1870 1880 1890
 1862 - - - - - CA-TTGG - - - - - 4507170-SPARC.seq
 2427 - - - - - AGCAGT - - - - - AGTG-TTTAA - 213614-QR-1.seq
 2501 - - - - - AAAAAT - - - - - GTACA-ATTAA - 21359870-Havin.seq
 4163 ACTTTGGAAGAACTTCTTAATGAACAGTTTATTTCTTATTC 15451924-SPOCK.seq
 1246 - - - - - AGAGGC - - - - - AGGT - CTGA - 13242264-FRE.seq

| | 1870 | 1880 | 1890 | |
|------|---|---|-------------------|--|
| 1869 | -A F G A - | T T G T C A T A A G B T T T T F A G C A T G T T C C | 4507170-SBARC.ssq | |
| 2442 | -G - - - A G - - - - - C H - - - - - C T A A G T A G C G C A T A C T | 213616-DR-1.ssq | | |
| 2517 | -G F A A G - - - - - - - F C A C A X A T G C R A P A T A C T G T A O A T E | 21350870-Revin.ssq | | |
| 4201 | C A T G A A C A G C G C T T T T A A T A G C A A T T F C C T F G B T A T C G A A C | 19451924-SPOCK.ssq | | |
| 1260 | A T G A A A - - - - - - - C C - | 13242264-FRR.ssq | | |

| | TTAAA--CAGAAATTTTCTTCAC | Mayoria |
|------|---|--------------------|
| -190 | | |
| | 1910 | |
| 1899 | TC--CCTTCTTCAC | 4507170-SFARO.seq |
| 2463 | TAAGGTCTCTGC | 213614-OR-1.seq |
| 2549 | CTGAA-CAGARGTTAAATTCAT | 21359870-Kevin.seq |
| 4241 | TTAAATGCCAGAAATTTGGTTCCACCATGATTAGTTATTTG | 15451924-SPOCK.seq |
| 1269 | | 13242264-FBP.seq |

- - - - - A C G T A C T T T G T T C T C T G T A T T A - - - - - A G T T A Mayoria
 1920 1930
 1912 - - - - - C C T C C G G T T T G T T C T T C - T A T T A - - - - - A - T C A 4507170-SFAC.seq
 2483 - - - - - A C G T A C C A A T A A T C T C T - T A G T A - - - - - A G - - A 213614-QR-1.seq
 2570 - - - - - A - G T A A T T T C A C T C T C T G C A T T G - - - - - A C T A 21359670-Kevin.seq
 4281 C C A T A G A A T G T A T C T T G T T C T T A C C A A T A T T C T A G C C A 1545194-SPOCK.seq
 1259 - - - - - T T T C T C G G G G T C - - - - - A G C T G 13242264-FRP.seq

| TGAGAA-ACCTTATGATTGAGTGT | | Mayoria |
|--------------------------------|------|--------------------|
| 1840 | 1950 | |
| 1930 AGAGAA-ACCTCAAGCTTAA--TGC | | 4507170-SHARC.seq |
| 2506 CGACTA-ACCTGATGACTGAGTGT | | 213614-GR-1.seq |
| 2597 TGAGAT-AATTAATGATTAACAT | | 21359870-Havin.seq |
| 4321 TGAGAA-TACCTAATACTGAGT | | 15451924-SPOCK.seq |
| 1284 TGAG--T--GAT | | 13242264-FRE.seq |

-----TCACCAATAG-GA Mayoria

-----360

1960 -----GATGCTCG-GA 4507170-SFARC.seq

2532 -----TCACAA----- 213614-QR-1.seq

2621 -----TAATGATAA-AA 21359870-Havin.seq

4361 TTCTCTCTCTCTTACGCGCATTTTCCTACCCACAGCAGCA 15451824-BROCK.seq

1291 -----ACACACAG-GA 18242264-FRZ.seq

AT -- AACGCATCTCTT -- AGA -- -- -- GTTCAT - C Mayoria

97 1980 1990

1970 - T -- CTCACAGGCGG -- AGA -- -- -- ACTCGGTC 4507170-SPARC.seq

2539 - - - AACCCCTTCGGT -- AGA -- -- -- ATT - - - A 213614-QR-1.seq

2632 AT - - - AATGCATTTTGT -- ATT -- -- -- GTTCATAA 21359970-Bavin.seq

4401 ATGACACGCTGTCTCTTCCAGCTCTGTCCTCATTC 15451526-BPOCK.seq

1300 A - - - - - - - - - - - - - - - - - - - 13242266-SPF.seq

Mayoria

TAGCATGT--GCATTT--
200
1993 ACGTCGAA--GCATTC-- 4507170-SPARC.seq
2356 TAGCATGT--GCATTT-- 213614-QR-1.seq
2656 TAGCATGT--GCATTC-- 21359870-Hevin.seq
4441 CAGCATGCCACTCAGCATTTTCTCTTCTTGTACCTTG 15481924-SPQR.seq
1301 AAGCATGT--GATGTT-- 13242264-X23.seq

```

-----ATGAGAAAC-----TUGA-AGCT-AGT----- Mayoria
2010 2020
2008 -----ATGAAAAA-----GCT-GCT----- 4507170-SPARC.seq
2571 -----ATAATACAC-----USATAAAA-ACT----- 213614-QR-1.seq
2671 -----A-AGAAAAA-----GGAATGCT-ACT----- 21359870-Havin.seq
4481 ACCGACGCTGTACBAGTAACATCCAGAGGCCCATCTACA 15451924-SPOCK.seq
1315 -----AGGGGGAAC-----ACT----- 13242264-FRP.seq

-----GTTTTGTTTATTATGTATATT-TTTTCAATAT Mayoria
2030 2040 2050
2022 -----TCTTATTAAATGATACAAACTCTCAGCATGAA 4507170-SPARC.seq
2592 -----ACTTTGAAAT-----AGGTTTTCT-TTTCCTGTTC 213614-QR-1.seq
2690 -----CTTTTGCTGCTTACGTCGATTTATTTTCAATAT 21359870-Havin.seq
4521 AGTGGCTGGTTTTGCTCTTTTATTAACCTTTTCTCAAAGT 15451924-SPOCK.seq
1327 -----GAGCTG----- 13242264-FRP.seq

G-TTAAATCT-----CCTGATAAA----- Mayoria
206
2053 G-TGAGAGG-----TTTCACAAA----- 4507170-SPARC.seq
2620 T-TTACTGT-----C-AGTA----- 213614-QR-1.seq
2722 C-TTAAATAC-----CCTAATAA----- 21359870-Havin.seq
4561 CACTGATGTTTGTCTCTTAAATGTATAGCAATTGTAAATG 15451924-SPOCK.seq
1333 -----GC-----CCTGCTGA----- 13242264-FRP.seq

-----GAGTCCATCAAAATCT----- Mayoria
2070 208
2070 -----TCTTTTAAAAAT----- 4507170-SPARC.seq
2633 -----GCTC-----TCT----- 213614-QR-1.seq
2730 -----GAGTCCATCAAAATCT----- 21359870-Havin.seq
4601 AGAGCCCATCAAAATCTGAGGTGTCAGTTTGTGTCCTAT 15451924-SPOCK.seq
1344 -----GG-----AAAT----- 13242264-FRP.seq

-----AAATAGTAATGTCAAAAAC----- Mayoria
2090 2100
2081 -----AAAAAGTAATGACTTAGAAR----- 4507170-SPARC.seq
2640 -----GCATAGAAATGTCAAAATAAAC----- 213614-QR-1.seq
2755 -----AAAAAAAABAAAAA----- 21359870-Havin.seq
4541 GTAGATGAAATAGTGTATGTAGCAAAAACCTAGTAATTC 15451924-SPOCK.seq
1350 -----AGG----- 13242264-FRP.seq

-----TGAAGTTATTTGGATA-AAA----- Mayoria
2110 2120
2102 -----TGAAGTTATTTGGATA-AAA----- 4507170-SPARC.seq
2661 -----AGATCTTGTCTTGTCTTCTC----- 213614-QR-1.seq
2776 -----AAAAAAAABAAAAA----- 21359870-Havin.seq
4681 TGAATGCTTTTCTCAGCTAGATTTATCTGCAATGTCAACAC 15451924-SPOCK.seq
1354 -----GAGCTTGTCTGCTG----- 13242264-FRP.seq

-----AAAAAAAAB----- Mayoria
213
2122 -----AAAAAAAAB----- 4507170-SPARC.seq
2678 -----AAAAAAAAB----- 213614-QR-1.seq
2797 -----AAAAAAAAB----- 21359870-Havin.seq
4721 AACTCTTTGGTTAAATAGTAATGCTTAACTGTAGTCTCTGA 15451924-SPOCK.seq
1368 ----- 13242264-FRP.seq

```

```

-----AAAXXXXXXXXXXXXX Mayoria
[REDACTED]
2131 -----AAA
2678 -----
2806 -----AAA
4761 G T A G G T G C A T T T G T C T G T C T C A A T A A T T T T A C T T T G T 1507170-SEARC.seq
1368 -----AGG 213614-GR-1.seq
                                     21359870-Revln.seq
                                     15451924-SPOCK.seq
                                     13242264-TRP.seq

                                     Mayoria

2133                                     4507170-SEARC.seq
2678                                     213614-GR-1.seq
2808                                     21359870-Revln.seq
4901 C T G C                                     15451924-SPOCK.seq
1370                                     13242264-TRP.seq

```

SEC ID N°

44
45
46
47
48

-----Mayoria

10 20 30

1 MK---TSEFFLCCLGTAARLEFNAR---LLSDHESKPTAETV SPARC.pro
1 MS---NLLI---TESTICAN.pro
1 MK---TGPFRLCLLGTAAALPTNAR---LLSDHESKPTAETV HEVIN.pro
1 MDEETRLST---A FRP.pro
1 MK---TVLLIICCLGSAFTTPTDPLNTQFGANGQXTARK QRL.pro

REFNTRIP-----EEKKST-----AVSTED-DSEHKAE Mayoria

40 50 60

36 APDNTAIFSLWAAARERERKET-----AVSTED-DSEHKAE SPARC.pro
8 AFEDVT-----ETQAC-----TDKEIA-----TESTICAN.pro
36 APDNTAIFSLWAAARERERKET-----AVSTED-DSEHKAE HEVIN.pro
11 LFCLETP-----TTRCDLVIFSGGEND---TRP.pro
38 KYTRSHMP-----ESENTG-----FVNGQDVLGHRAT QRL.pro

KSSVLLAA-----EESH---DASAEQCKSS-----KQ Mayoria

70 80 90 100

70 KASVLEKSK-----EESH---EQSAEQCKSSQELGLKDO SPARC.pro
25 ---NLASRLKDFWFGALRE---DANKVIRPTA---SOT TESTICAN.pro
70 KASVLEKSK-----EESH---EQSAEQCKSSQELGLKDO HEVIN.pro
33 ---ELAS-----GTFQ---PFVTHLKSVS---DRI FRP.pro
66 KAEVPVLD-----TQKDEFWASHAKQGGCD---GEN QRL.pro

EDGCGHLSVNLLYAF-SGGGLDIREDNIEP-QEKRLSEET Mayoria

110 120 130

101 EDSGDGHLGVNLEYAF-TBGTLDIKEDMIEP-QEKRLSEET SPARC.pro
53 TQGRF---DTSILFTCKDGLGWHFNEZPMN-YDLLELDHSE TESTICAN.pro
101 EDSGDGHLGVNLEYAF-TBGTLDIKEDMIEP-QEKRLSEET HEVIN.pro
54 SKGGY---EVLZRES-SVOVPE-PTKYTLGRFVREVTTFE FRP.pro
93 QTKNSLRASINFLTLH-SNFGLABDQWESNS-GSARLQSS QRL.pro

D-----FLA- Mayoria

140

139 D-----FLAP SPARC.pro
89 INAXYLDKYRFOCIKPLENSCDSFKDGKLSNNWNCYCFQ---TESTICAN.pro
139 D-----FLAP HEVIN.pro
89 V-----LE---FRP.pro
132 E-----HQ- QRL.pro

-----TTDANQQESIT-----KRENNQSQPRNTYSHHQL-NRS Mayoria

120 150 170

144 GVSSTTDSNQQESIT-----KRENNQSQPRNTYSHHQL-NRS SPARC.pro
127 ---KRENNQSQPRNTYSHHQL-NRS TESTICAN.pro
144 GVSSTTDSNQQESIT-----KRENNQSQPRNTYSHHQL-NRS HEVIN.pro
82 ---FRP.pro
135 ---PRRRKRENNMAQDVALROESVDPALGLVARNNTWXY QRL.pro

SKHSQG---LREQGNIQO---YSHGEEEEKKEPGEVGTEND Mayoria

180 190 200 210

179 SKHSQG---LRDQGNQEQDFNISHGEEEEKKEPGEVGTEND SPARC.pro
127 -KPGQLPQGNMNRICK---LSKQ---TESTICAN.pro
179 SKHSQG---LRDQGNQEQDFNISHGEEEEKKEPGEVGTEND HEVIN.pro
82 -RSVTL---EKETEPIEP---STQA---FRP.pro
171 NKNTVG---LDENNGSE---EEHAGEEEDDEWGEETDYRD QRL.pro

220 230

217 NQZ --- RKTFLPRRHA --- NS SPARC.pro
 147 --- KSLDGAFTIRCHNEEGYKATQCHGSEFCQCVCKYGN TESTICAN.pro
 217 NQZ --- RKTFLPRRHA --- NS HEVIN.pro
 110 --- NAAATAGEAE --- UN FRP.pro
 206 MKH --- RARGTSHGRE --- YR QRI.pro

---SQSE---SNTQSDIIEESDQPTQVSKMQEDEFDQGNEE Mayoria

240 250 260

232 -KQEE---DNTQSDDIIEESDQPTQVSKMQEDEFDQGNQE SPARC.pro
 184 -FLAG---SRKQCTV---SCSE TESTICAN.pro
 232 -KQEE---DNTQSDDIIEESDQPTQVSKMQEDEFDQGNQE HEVIN.pro
 122 -ELGGTWTTSQESTAL---SRTK FRP.pro
 221 RWQNE---NRRPSSGZFLRDSLSLPVRIKRRHGKFSMNEES QRI.pro

Q---EDNSNARMMEESNASLV---N Mayoria

270 280

269 Q---EDNSNARMMEESNASLV---N SPARC.pro
 199 E---QETSQDFTGSSGSSVVL---TESTICAN.pro
 269 Q---EDNSNARMMEESNASLV---N HEVIN.pro
 140 G---EDQDFTVEENKSVGL---FRP.pro
 258 QKLYKEGKLELSKKNNMEOQGEKROSEASKEHFOVVNR QRI.pro

KHIQETEFWQSQEGKTGLEATSNHKEETEERTV---SEAL Mayoria

290 300 310 320

288 KHIQETEFWQSQEGKTGLEATSNHKEETEERTV---SEAL SPARC.pro
 218 ---LDDLODERHSLGPKD---TESTICAN.pro
 288 KHIQETEFWQSQEGKTGLEATSNHKEETEERTV---SEAL HEVIN.pro
 156 ---QVLENNKKAALCKE---FRP.pro
 296 KERA VTKRQDKEGSNAREDDNDSDGSEEDLGNVWAEAV QRI.pro

---ETDDGNTTFRNAGVDDDDGDDDDGDDGDTDGFRHSAED--- Mayoria

330 340 350

322 MEPTDDGNTTFRNAGVDDDDGDDDDGDDGDTDGFRHSAED SPARC.pro
 229 ---KEGKLRVETRAVREDDDDDDDDDDKDDDEVLCLFIGHTFL TESTICAN.pro
 322 MEPTDDGNTTFRNAGVDDDDGDDDDGDDGDTDGFRHSAED HEVIN.pro
 170 ---QA---MAYARALVVGFEID---YHDDL FRP.pro
 337 ---YEMMERMQSNDQDSITNKQKEEITAGDDSGVYREMQ QRI.pro

---D---I---FG---QAFLEAERAQ---SIAYRL---KIEBQR Mayoria

360 370 380

360 ---DYFI---PS---QAFLEAERAQ---SIAYRL---KIEBQR SPARC.pro
 265 LKIMSPIL---GFPFGHSEFQDKHSSR---LSTALI---SCLECK TESTICAN.pro
 360 ---DYFI---PS---QAFLEAERAQ---SIAYRL---KIEBQR HEVIN.pro
 191 ---F---STAD---APGASRLHACVNF---VD---LCKRAN FRP.pro
 373 ---D---Y---KG---DKIKDVTRESE---DNKXKREPERENSSX QRI.pro

EXVQENRV---IGTTERGERQZAKKAENSSNE Mayoria

390 400 410

388 EXVQENRV---IGTTERGERQZAKKAENSSNE SPARC.pro
 301 E---QNAVSSGWKEVANTHEAMAS---TESTICAN.pro
 388 EXVQENRV---IGTTERGERQZAKKAENSSNE HEVIN.pro
 218 E---DMMWV---DQITA---VESMNSTENHDDVETTEGSEYR QRI.pro
 401 QQLQTSSE---VESMNSTENHDDVETTEGSEYR

Mayoria

420 430 440 450

417 EETSSSEGNMRLVHAYDSCMSFQCKRGHICQKADQQQKPRHCVC SPARC.pro
 323 -----VNKII-DK TESTICAN.pro
 417 EETSSSEGNMRLVHAYDSCMSFQCKRGHICQKADQQQKPRHCVC HEVIN.pro
 223 -----MQAFR-RF FRF.pro
 430 EESARNSTGKAL-PDLCHNFHCKRAGKVCQADKQCKPSCIC QRI.pro

ODPVFCFFTEKFLDQVCSTDNQTYASSCHLEFATKCRLEGTK Mayoria

460 470 480 490

457 QDPVTCFFTEKFLDQVCSTDNQTYASSCHLEFATKCRLEGTK SPARC.pro
 329 TESTICAN.pro
 457 QDPVTCFFTEKFLDQVCSTDNQTYASSCHLEFATKCRLEGTK HEVIN.pro
 235 FRF.pro
 459 QDEAACPFSTKDYKAVCGTDNNTYDGTICOLEFGTKQOLEGTX QRI.pro

KGRQLQLDYFGACKSIFPTCTDYEVIQFPLMRDNLKNILM Mayoria

500 510 520 530

497 KGRQLQLDYFGACKSIFPTCTDYEVIQFPLMRDNLKNILM SPARC.pro
 329 TESTICAN.pro
 497 KGRQLQLDYFGACKSIFPTCTDYEVIQFPLMRDNLKNILM HEVIN.pro
 235 FRF.pro
 509 MGRQLHLDYNGACKRIPTCTDYEVIQFPLMRDNLKNILM QRI.pro

QLYEANSSEHAGYLNKQANKVKKIYLDKRLLAGDHPIDL Mayoria

540 550 560 570

537 QLYEANSSEHAGYLNKQANKVKKIYLDKRLLAGDHPIDL SPARC.pro
 329 TESTICAN.pro
 537 QLYEANSSEHAGYLNKQANKVKKIYLDKRLLAGDHPIDL HEVIN.pro
 235 FRF.pro
 549 QYXERDQDTSALTEKQANKVKKIYLDKRLVSGERFVSL QRI.pro

LIRDFFKNYMYVYFVHWQFSELDOHPMDAYLTSESLAPL Mayoria

580 590 600 610

577 LIRDFFKNYMYVYFVHWQFSELDOHPMDAYLTSESLAPL SPARC.pro
 329 TESTICAN.pro
 577 LIRDFFKNYMYVYFVHWQFSELDOHPMDAYLTSESLAPL HEVIN.pro
 235 FRF.pro
 589 LIRDFFKNYMYVYFVHWQFSELDOHPMDAYLTSESLAPL QRI.pro

RASLVEMEHOCITAFEECDENKDKHITLKEWGHCFGIKEE Mayoria

620 630 640 650

617 RASLVEMEHOCITAFEECDENKDKHITLKEWGHCFGIKEE SPARC.pro
 329 TESTICAN.pro
 617 RASLVEMEHOCITAFEECDENKDKHITLKEWGHCFGIKEE HEVIN.pro
 235 FRF.pro
 629 RASLVEMEHOCITAFEECDENKDKHITLKEWGHCFGIKEE QRI.pro

DIDENLLF Mayoria

660

657 DIDENLLF SPARC.pro
 329 TESTICAN.pro
 657 DIDENLLF HEVIN.pro
 235 FRF.pro
 669 DINENLLF QRI.pro

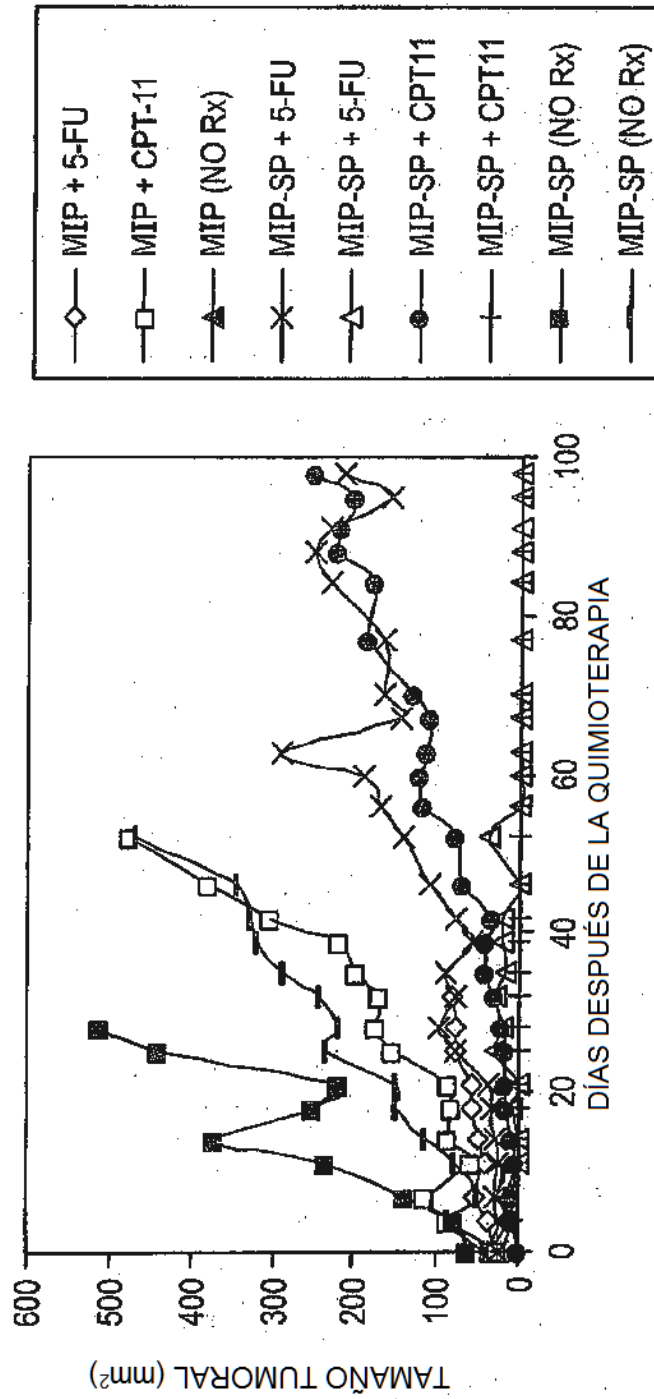


FIG. 14

Figura 15

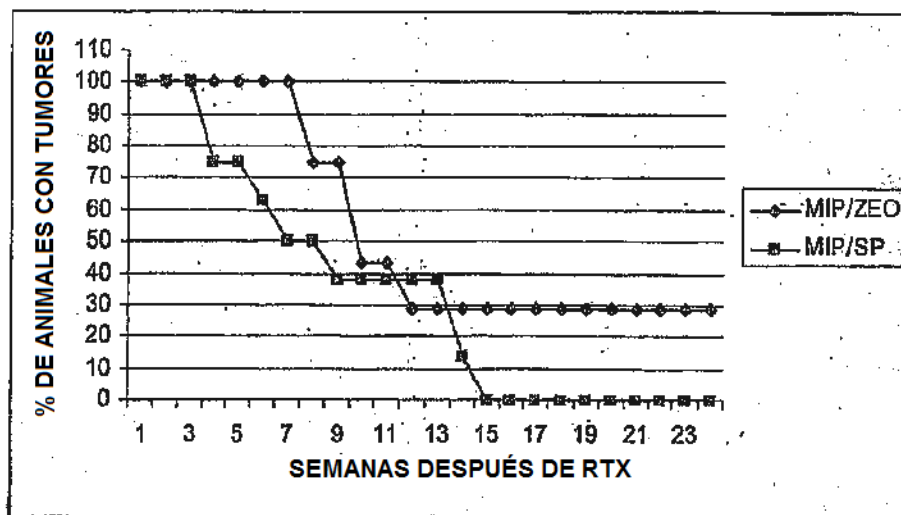


Figura 16

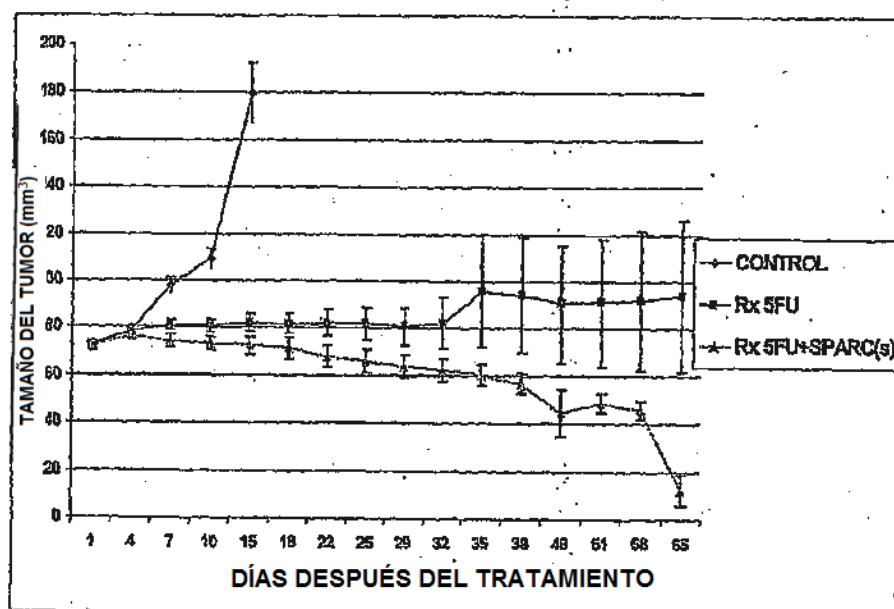


Figura 17

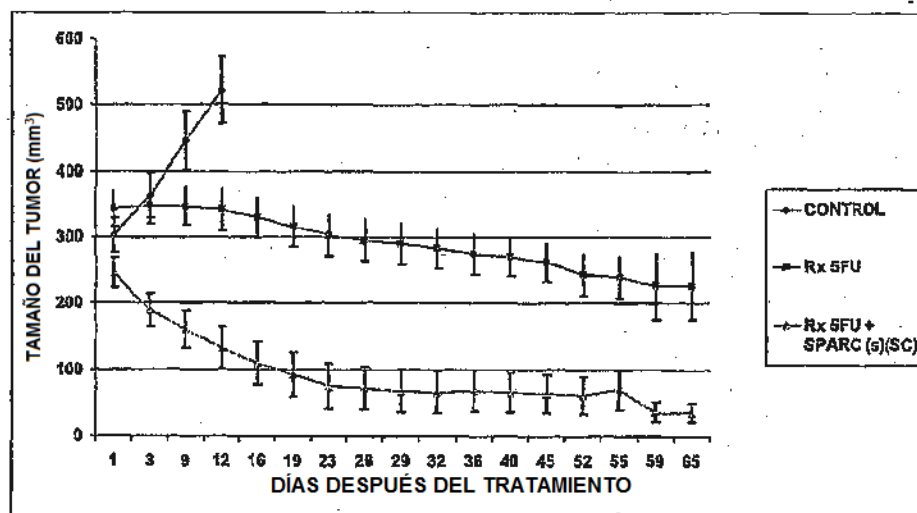


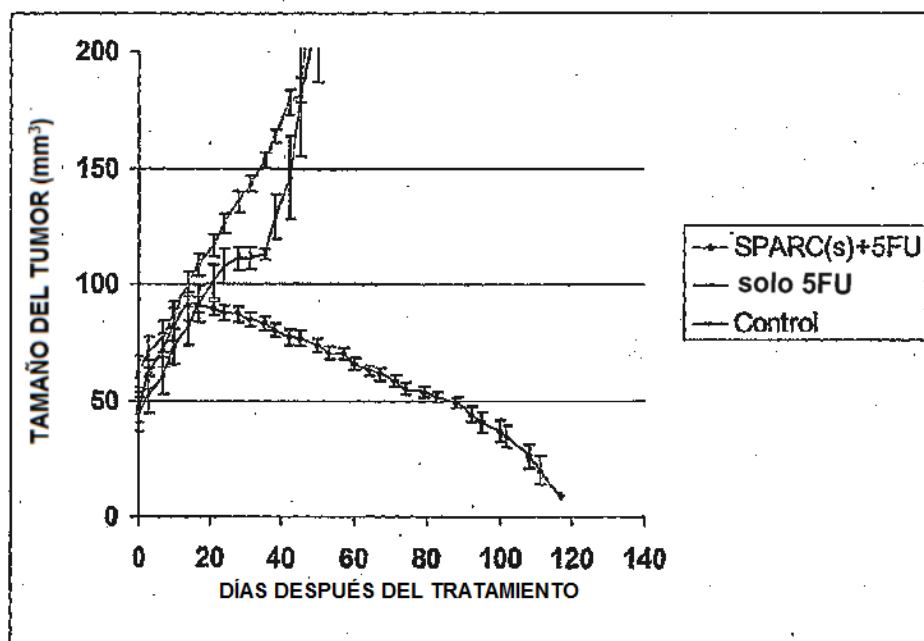
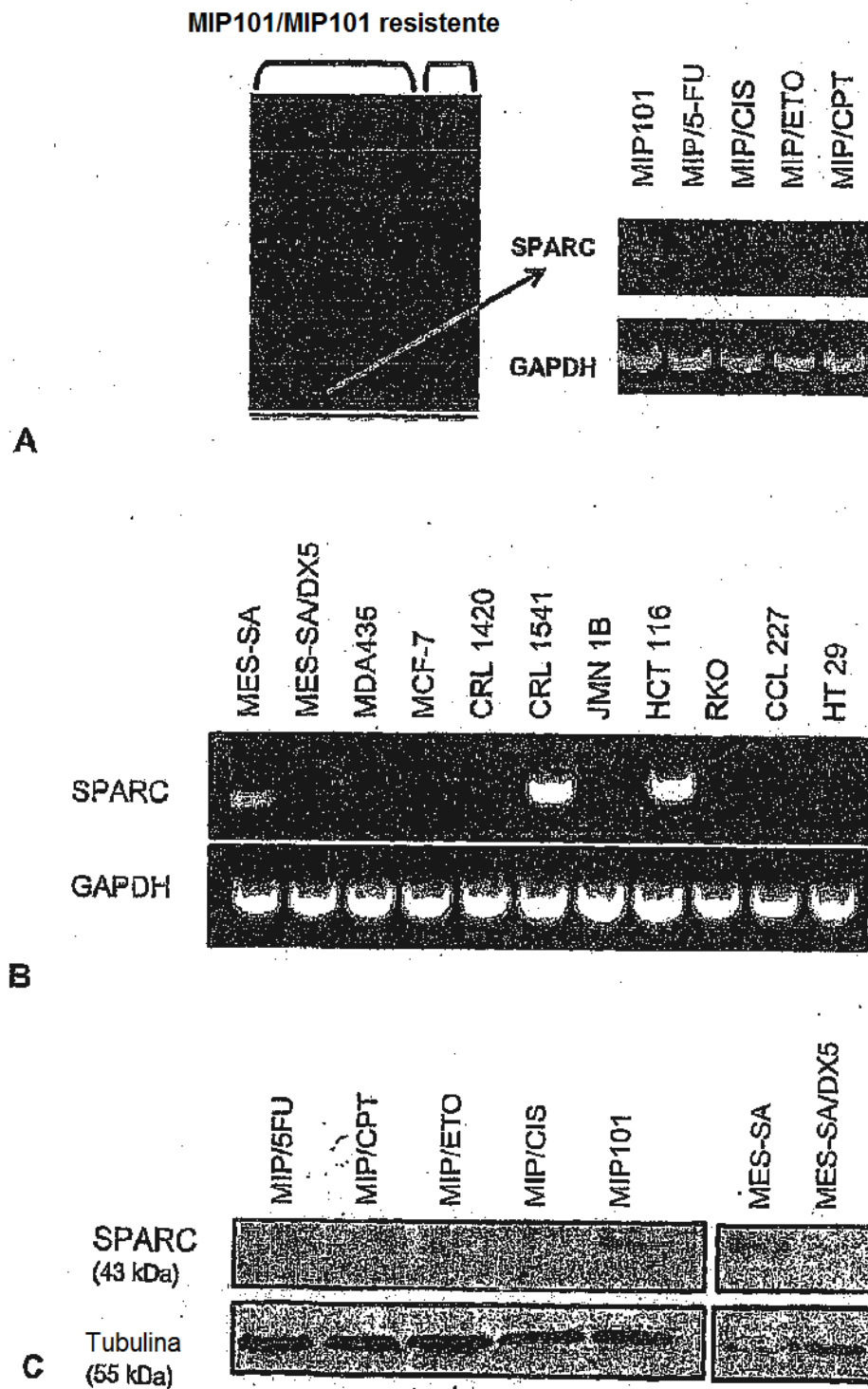
Figura 18

Figura 19



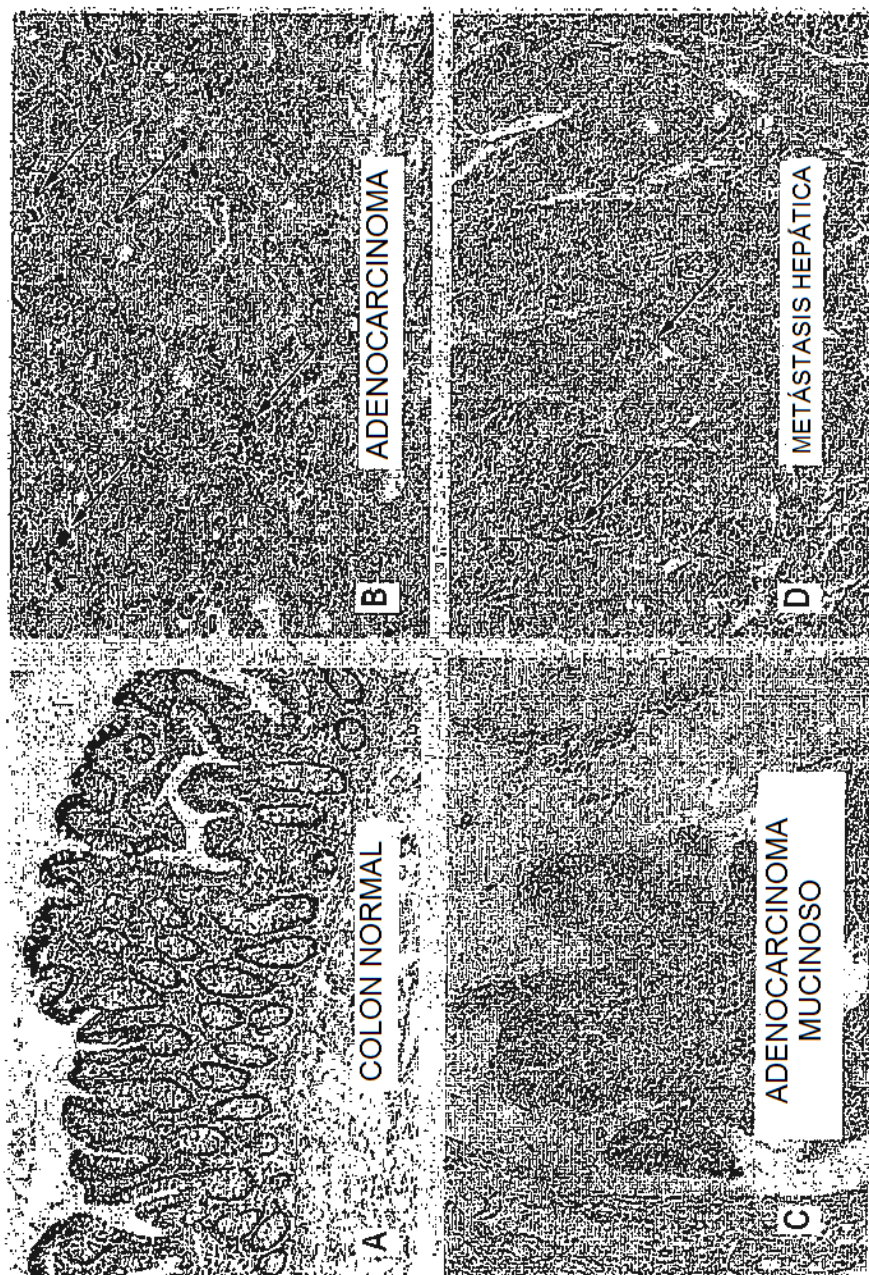


FIG. 20

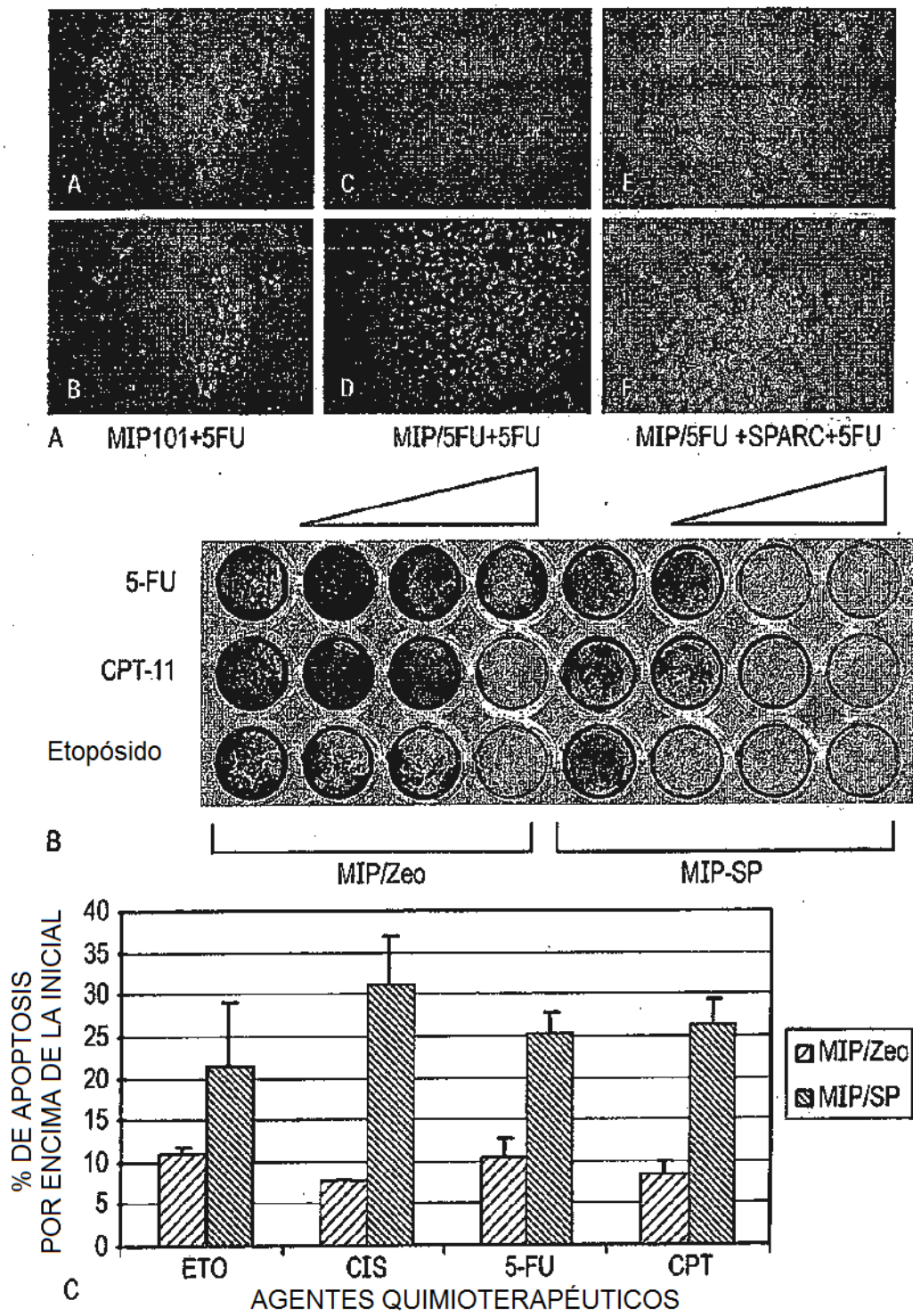


FIG. 21