



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 268 886**

51 Int. Cl.:

A61K 48/00 (2006.01)

C07K 14/705 (2006.01)

A61K 38/17 (2006.01)

G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **99956430 .5**

86 Fecha de presentación : **06.10.1999**

87 Número de publicación de la solicitud: **1121157**

87 Fecha de publicación de la solicitud: **08.08.2001**

54 Título: **Un nuevo componente en la vía de señalización de hedgehog.**

30 Prioridad: **06.10.1998 SE 9803393**

45 Fecha de publicación de la mención BOPI:
16.03.2007

45 Fecha de la publicación del folleto de la patente:
16.03.2007

73 Titular/es: **Karolinska Innovations AB.
171 77 Stockholm, SE
Pharmacia & Upjohn Company L.L.C.**

72 Inventor/es: **Zaphiropoulos, Peter, G.;
Unden, Anne, Birgitte;
Toftgard, Rune;
Rahnama, Fahimeh y
Hollingsworth, Robert E.**

74 Agente: **Ungría López, Javier**

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

Un nuevo componente en la vía de señalización de hedgehog.

5 **Campo técnico**

La presente invención se refiere a nuevas moléculas, tales como proteínas, polipéptidos y nucleótidos, implicados en la vía de señalización de hedgehog con una supuesta implicación en el desarrollo embrionario y carcinogénesis. La invención también se refiere a diversos usos nuevos y ventajosos de las moléculas de acuerdo con la invención, por ejemplo, en el diagnóstico y terapia.

Antecedentes

En el estudio del desarrollo de las células, la mosca de la fruta se ha usado ampliamente como modelo, ya que es menos compleja que las células de mamífero.

La formación de patrones tiene lugar a través de una serie de etapas lógicas, reiteradas muchas veces durante el desarrollo de un organismo. Visto desde una perspectiva evolutiva más amplia, entre las especies, se observa el mismo tipo de formaciones de patrones reiterativos. Se ha descrito el dogma central de la formación de patrones (Lawrence y Struhl, 1996). Se definen tres etapas interbloqueantes y solapantes. En primer lugar, la información posicional en forma de gradientes morfogénicos sitúa las células en etapas no solapantes, encontrando cada conjunto un compartimiento. En segundo lugar, cada uno de estos compartimientos adquiere una dirección genética, como resultado de la función de genes “selectores” activos, que especifican el destino celular en un compartimiento y también muestran a las células y sus descendientes cómo comunicarse con células en los compartimientos adyacentes. La tercera etapa implica interacciones entre células en compartimientos adyacentes, iniciando nuevos gradientes morfogénicos, que organizan directamente el patrón.

Tomando estas etapas con mayor detalle, se descubre que la primera etapa en la formación de patrones es la definición de conjuntos de células en cada primordio. Las células se sitúan de acuerdo con sus posiciones con respecto a los ejes dorsoventral y anterior/posterior por gradientes morfogénicos. La asignación de las células en el eje dorsoventral constituye las capas germinales, tales como el mesodermo y el neuroectodermo.

En la segmentación, la segunda etapa (la especificación del destino celular en cada compartimiento) se realiza pero por el gen *engrailed* y los elementos del complejo bitórax. *Engrailed* define los compartimientos anterior y posterior tanto en la segmentación como en la especificación de extremidades.

La tercera etapa en la formación de patrones, la secreción de morfogenes, funciona diferenciando los patrones en los compartimientos (y estableciendo de este modo la polaridad de los segmentos). Inicialmente, todas las células en un compartimiento son equipolentes, pero llegan a diversificarse para formar el patrón. La formación de patrones depende de los gradientes de morfogenes, gradientes iniciados a lo largo de los límites del compartimiento. Dichos gradientes se establecen por una señal de corto alcance en todas las células del compartimiento en el que es activo el gen selector *engrailed* mencionado anteriormente. Para la polaridad de los segmentos, esta señal es Hedgehog. En el compartimiento adyacente, el gen selector está inactivo, asegurando que las células sean sensibles a la señal. El intervalo de señal de Hedgehog es probablemente sólo de unas pocas filas de células; las células respondedoras llegan a formar una fuente lineal de un morfógeno de largo alcance, que se difunde hacia el exterior en todas las direcciones. Existen tres Hedgehog conocidas, Sonic (SHH), Indian (IHH) y Desert (DHH). Las proteínas que codifican pueden sustituirse entre sí, pero en animales de tipo silvestre, sus distintas distribuciones provocan actividades únicas. SHH controla la polaridad del crecimiento de las extremidades, dirige el desarrollo de las neuronas en el tubo neural ventral y forma patrones de somitos. IHH controla el desarrollo del hueso endocondral y DHH es necesaria para la espermiogénesis. Los genes de hedgehog de vertebrados se expresan en muchos tejidos diferentes, incluyendo el sistema nervioso periférico, el cerebro, los pulmones, el hígado, los riñones, los primordios dentales, los genitales y el endodermo del intestino posterior y el intestino anterior.

Por tanto, se han identificado los genes de polaridad de los segmentos en las moscas como mutaciones, que cambian el patrón de las estructuras de los segmentos corporales. Mutaciones en estos genes provocan que los animales desarrollen los patrones cambiados en las superficies de los segmentos corporales, afectando los cambios al patrón a lo largo del eje de la cabeza a la cola. Por ejemplo, mutaciones en el gen *patched* provoca que cada segmento corporal se desarrolle sin las estructuras normales en el centro de cada segmento. En su lugar, hay una imagen especular del patrón normalmente encontrado en el segmento anterior. Por tanto, las células en el centro del segmento componen las estructuras erróneas, y las dirigen en la dirección incorrecta con referencia a la polaridad de la cabeza a la cola global del animal.

Se conocen aproximadamente dieciséis genes de esta clase. Las proteínas codificadas incluyen quinasas, factores de transcripción, una proteína de unión celular, dos proteínas secretadas llamadas wingless (WG) y la Hedgehog mencionada anteriormente (HH), una proteína transmembrana única llamada patched (PTC) y algunas proteínas nuevas no relacionadas con ninguna proteína conocida. Se cree que todas estas proteínas trabajan juntas en las vías de señalización que informan a las células acerca de sus alrededores para establecer los destinos celulares y polaridades.

Se ha propuesto PTC como un receptor para la proteína HH en base a experimentos genéticos en moscas. Un modelo para la relación es que PTC funciona a través de una vía muy desconocida para inactivar su propia transcripción y la transcripción del gen de polaridad de segmentos *wingless*. Este modelo propone que la proteína HH, secretada por células adyacentes, se une al receptor PTC, lo inactiva y de este modo evita que PTC anule su propia transcripción y la de *wingless*. Varios experimentos han demostrado acontecimientos coordinados entre PTC y HH.

El gen *patched* humano (PTCH) se identificó recientemente como el gen responsable del síndrome de carcinoma de células basales nevoides (NBCCS), también conocido como el síndrome de Gorlin, que es un trastorno dominante autosómico que predispone a cáncer y a defectos en el desarrollo (Gorlin (1995) *Dermatologic Clinics* 13:113-125) caracterizado por múltiples carcinomas de células basales (BCC), meduloblastomas y fibromas de ovario así como numerosas anormalidades en el desarrollo (Hahn, H., Wicking, C., Zaphiropoulos, P.G., Gailani, M.R., Shanley, S., Chidambaram, A., Vorechovsky, I., Holmberg, E., Undén, A.B., Gillies, S., Negus, K., Smyth, I., Pressman, C., Leffell, D.J., Gerrard, B., Goldstein, A.M., Dean, M., Toftgard, R., Chenevix-Trench, G., Wainright, B. y Bale, A.E. (1996): "Mutations of the human homolog of *Drosophila* patched in the nevoid basal cell carcinoma syndrome", *Cell* 85, 841-851; y Jonson, R.L., Rothman, A.L., Xie, J., Goodrich, L.V., Bare, J.W., Bonifas, J.M., Quinn, A.G., Myers, R.M., Cox, D.R., Epstein, E.H. Jr y Scout, M.P. (1996): "Human homolog of patched a candidate gene for the basal cell nevus syndrome", *Science* 272, 1668-1671). PTCH codifica un receptor de membrana del dominio amino terminal escindido autolíticamente (proteína con corte y empalme) de *sonic hedgehog* (SHH) (Mariano, V., Davey, R.A., Zuo, Y., Cunningham, J.M. y Tabin, C.J. (1996): "Biochemical evidence that patched is the Hedgehog receptor", *Nature* 384, 176-179; y Stone, D.M., Hynes, M., Armanini, M., Swanson, T.A., Gu, Q., Johnson, R.L., Scott, M.P., Pennica, D., Goddard, A., Phillips, H., Noll, M., Hooper, J.E. de Sauvage, F. y Rosenthal, A (1996): "The tumor-suppressor gene patched encodes a candidate receptor for Sonic hedgehog", *Nature* 384, 129-134). En el estado de no-señalización, se cree que PTCH inhibe la señalización consecutiva de otras proteínas de membrana, *smoothened* (SMO), sin embargo, la unión de SHH a PTCH alivia esta inhibición (Goodrich, L.V., Milenkovic, L., Higgins, K.M. y Scott, M.P. (1997): "Altered neural cell fates and medulloblastoma in mouse patched mutants", *Science* 277, 1109-1113). Esta cascada de acontecimientos de señalización, mejor caracteriza in *Drosophila*, también implica varios componentes intracelulares incluyendo *fused* (una serina treonina quinasa), el supresor de *fused*, *costal 2*, y *cubitus interruptus* (Ruiz i Altaba, A.: "Catching a Glimpse of Hedgehog" (1997) *Cell* 90, 193-196). El último es un factor de transcripción que regula de forma positiva la expresión de genes diana que también incluyen PTCH.

Se han identificado mutaciones en el gen de PTCH en BCC esporádicos y familiares (Gailani, M.R., Stahle-Bäckdahl, M., Leffell, D.J., Glynn, M., Zaphiropoulos, P.G., Pressman, C., Undén, A.B., Dean, M., Brash, D.E., Bale, A.E. y Toftgard, R. (1996): "The role of human homologue of *Drosophila* patched in sporadic basal cell carcinomas" *Nature Genet.* 14, 78-81). La ausencia de la proteína PTCH normal en estas células permite que suceda la señalización constitutiva de SMO, provocando la acumulación de ARNm de PTCH mutantes (Undén, B.A., Zaphiropoulos, P.G., Bruce, K., Toftgard, R., y Stahle-Bäckdahl, M. (1997): "Human patched (PTCH) mRNA is overexpressed consistently in tumor cells of both familial and sporadic basal cell carcinoma", *Cancer Res.* 57, 2336-2340).

El documento WO 96/11260 describe el aislamiento de genes *patched* y el uso de la proteína PTC para identificar ligandos, diferentes al ligando Hedgehog establecido, que se unan al mismo.

Sin embargo, aún existe la necesidad de comprender adicionalmente la señalización celular de SHH/PTCH, que puede proporcionarse por la descripción de genes adicionales, péptidos y proteínas implicados en la misma.

Sumario de la invención

La presente invención proporciona una etapa significativa dirigida al entendimiento de la vía descrita anteriormente. Se ha clonado y secuenciado un nuevo gen tipo *patched* humano (PTCH2) por una combinación de análisis de bibliotecas de ADNc y RACE. Se han identificado varias formas de ARNm con un corte y empalme alternativo de PTCH2, incluyendo transcritos que carecen de segmentos que se cree que están implicados en la unión de sonic hedgehog (SHH) y ARNm con exones en el extremo 3' terminal definidos de forma diferente. Por consiguiente, la invención se refiere a dos variantes de corte y empalme que producen diferentes extremos C-terminales como se describe en la Figura 2B.

Breve descripción de los dibujos

La Figura 1 muestra la secuencia genómica de la SEC ID N° 5, en la que los exones y los intrones están indicados en la secuencia genómica del nuevo gen de patched 2 humano.

La Figura 2A describe una comparación de secuencia de aminoácidos de las secuencias de PTCH2 (líneas superiores) y PTCH1 (líneas inferiores) humanas.

La Figura 2B es una representación de los acontecimientos de corte y empalme alternativos que producen diferentes extremos C-terminales.

La Figura 2C es una representación de las diferentes variaciones de transcritos que han sufrido corte y empalme que abarcan las secuencias del exón 1 y el exón 2.

La Figura 3A es una fotomicrografía de campo oscuro de un tumor BCC hibridado con una sonda antisentido marcada con ³⁵S que muestra una señal abundante de ARNm de PTCH1 (granos brillantes) en todas las células de tumor BCC.

La Figura 3B describe la sobreexpresión de ARNm de PTCH2 en BCC y se expresa principalmente en contraste en las células basaloideas en la periferia de las redes tumorales.

La Figura 3C es otro BCC que muestra una fuerte señal de ARNm de PTCH2 en la periferia de la red tumoral (Tu), mientras que no se detecta señal en la epidermis (Ep).

La Figura 3D son secciones del mismo tumor (C) hibridado con la sonda con sentido de PTCH2 que no mostró señal.

La Figura 3E muestra la inmunorreactividad para Ki-67.

La Figura 3F describe cómo las redes tumorales bajo un aumento de alta potencia muestran una señal de ARNm de PTCH2 abundante (granos negros) en las células tumorales basaloideas oscuras y una señal inferior en el centro (flecha).

Definiciones

Los términos “polipéptidos”, “péptido” y “proteína” se usan indistintamente en este documento para referirse a un polímero de restos de aminoácidos. Los términos se aplican a polímeros de aminoácidos en los que uno o más restos de aminoácidos son un análogo químico artificial de un aminoácido de origen natural correspondiente, así como a polímeros de aminoácidos de origen natural.

Las expresiones “aislado”, “purificado” o “biológicamente puro” se refieren a un material que está sustancial o esencialmente libre de componentes que lo acompañan normalmente como se encuentra en su estado nativo.

La expresión “ácido nucleico” se refiere a un polímero de desoxirribonucleótidos o ribonucleótidos en forma monocatenaria o bicatenaria, y salvo que se limite de otro modo, incluye análogos conocidos de nucleótidos naturales que pueden funcionar de un modo similar a los nucleótidos de origen natural.

Un “marcador” es una composición detectable por medios espectroscópicos, fotoquímicos, bioquímicos, inmunológicos, o químicos. Por ejemplo, los marcadores útiles incluyen ³²P, colorantes fluorescentes, reactivos electrodenso, enzimas (por ejemplo, como se usa habitualmente en un ELISA), biotina, dioxigenina, o haptenos y proteínas para los que están disponibles antisueros o anticuerpos monoclonales (por ejemplo, el péptido de la SEC ID N° 1 puede hacerse detectable, por ejemplo, incorporando un radio-marcador en el péptido, y puede usarse para detectar anticuerpos específicamente reactivos con el péptido).

Como se usa en este documento una “sonda de ácido nucleico” se define como un ácido nucleico capaz de unirse a un ácido nucleico diana de secuencia complementaria a través de uno o más tipos de enlaces químicos, habitualmente a través de apareamiento de bases complementarias, habitualmente a través de la formación de enlaces de hidrógeno. Como se usa en este documento, una sonda puede incluir bases naturales (es decir, A, G, C, o T) o modificadas (7-desazaguanosina, inosina, etc.). Además, las bases de una sonda pueden unirse por un enlace diferente al enlace fosfodiéster, siempre que no impida la hibridación. Por tanto, por ejemplo, las sondas pueden ser ácidos peptidonucleicos en los que las bases constituyentes están unidas por enlaces peptídicos en lugar de enlaces fosfodiéster. Un especialista en la técnica entenderá que las sondas pueden unirse a secuencias diana que carecen de la complementariedad completa con la secuencia de la sonda dependiendo de la rigurosidad de las condiciones de hibridación. Las sondas preferiblemente se marcan directamente como con isótopos, cromóforos, luminóforos, cromógenos, o pueden marcarse indirectamente tal como con biotina a la que puede unirse después un complejo de estreptavidina. Ensayando la presencia o ausencia de la sonda se puede detectar la presencia o ausencia de la secuencia o subsecuencia seleccionada.

Una “sonda de ácido nucleico marcada” es una sonda de ácido nucleico que está unida, covalentemente, a través de un enlace, o a través de enlaces iónicos, de van der Waals o de hidrógeno a un marcador de modo que puede detectarse la presencia de la sonda detectando la presencia del marcador unido a la sonda.

La expresión “ácido nucleico diana” se refiere a un ácido nucleico (a menudo obtenido de una muestra biológica), para el que se diseña una sonda de ácido nucleico que hibrida específicamente. Lo que debe detectarse es la presencia o ausencia del ácido nucleico diana, o la cantidad del ácido nucleico diana que debe cuantificarse. El ácido nucleico diana tiene una secuencia que es complementaria a la secuencia del ácido nucleico de la sonda correspondiente dirigida a la diana. La expresión ácido nucleico diana puede referirse a la subsecuencia específica de un ácido nucleico más grande al que se dirige la sonda o a la secuencia global (por ejemplo, gen o ARNm) cuyo nivel de expresión se desea detectar. La diferencia en el uso será evidente a partir del contexto.

El término “recombinante” cuando se usa con referencia a una célula, o ácido nucleico, o vector, indica que la célula, o ácido nucleico, o vector, se ha modificado por la introducción de un ácido nucleico heterólogo o la alteración de un ácido nucleico nativo, o que la célula se obtiene de una célula modificada de este modo.

El término “idéntico” en el contexto de dos secuencias de ácidos nucleicos o polipéptidos se refiere a los restos en las dos secuencias que son iguales cuando se alinean para una correspondencia máxima. El alineamiento óptimo de las secuencias para la comparación puede realizarse, por ejemplo, por el algoritmo de homología local de Smith y Waterman (1981) *Adv. Appl. Math.* 2: 482, por el algoritmo de alineamiento de homología de Needleman y Wunsch (1970) *J. Mol. Biol.* 48: 443, por la búsqueda del método de similitud de Pearson y Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, por herramientas informáticas de estos algoritmos (GAP, GESTFTT, FASTA, y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison WI) o por inspección. El algoritmo BLAST realiza un análisis estadístico de la similitud entre dos secuencias; véase, por ejemplo, Karlin y Altschul (1993) *Proc. Nat'l Acad Sci USA* 90: 5873-5787.

La expresión “identidad sustancial” o “similitud sustancial” en el contexto de un polipéptido indica que el polipéptido comprende una secuencia con una identidad de secuencia de al menos el 70% con una secuencia de referencia, o preferiblemente una identidad de secuencia del 80%, o más preferiblemente del 85% con la secuencia de referencia, o más preferible una identidad del 90% en una ventana de comparación de aproximadamente 10-20 restos de aminoácidos. Un indicio de que dos secuencias polipeptídicas sean sustancialmente idénticas es que un péptido sea inmunológicamente reactivo con anticuerpos producidos contra el segundo péptido. Por tanto, un polipéptido es sustancialmente idéntico a un segundo polipéptido, por ejemplo, cuando los dos péptidos difieren sólo en una sustitución conservativa.

Un indicio de que dos secuencias de ácido nucleico sean sustancialmente idénticas es que el polipéptido que codifica el primer ácido nucleico sea inmunológicamente reactivo de forma cruzada con el polipéptido codificado por el segundo ácido nucleico. Otro indicio de que dos secuencias de ácido nucleico sean sustancialmente idénticas es que las dos moléculas hibriden entre sí en condiciones rigurosas.

La expresión “que hibride específicamente con”, se refiere a la unión, formación de dúplex, o hibridación de una molécula sólo con una secuencia de nucleótidos particular en condiciones de rigurosidad cuando la secuencia está presente en una mezcla de ADN o ARN compleja (por ejemplo, celular total). La expresión “condiciones de rigurosidad” se refiere a condiciones en las que hibridará una sonda con su subsecuencia diana, pero no con otras secuencias. Las condiciones de rigurosidad dependen de la secuencia y serán diferentes en diferentes circunstancias. Secuencias más largas hibridan específicamente a temperaturas más altas. Generalmente, las condiciones de rigurosidad se seleccionan para que sean aproximadamente 5°C inferiores al punto de fusión térmica T_m para la secuencia específica a una fuerza iónica y pH definidos. La T_m es la temperatura (con una fuerza iónica, pH, y concentración de ácido nucleico definidos) a la que hibrida el 50% de las sondas complementarias a la secuencia diana con la secuencia diana en equilibrio. (Como la secuencia diana está generalmente presente en exceso, a T_m , el 50% de las sondas estarán ocupadas en equilibrio). Típicamente, las condiciones de rigurosidad serán aquellas en las que la concentración salina sea de iones Na de menos de aproximadamente 1,0 M, típicamente una concentración de iones Na de aproximadamente 0,01 a 1,0 M (u otras sales) a un pH de 7,0 a 8,3 y la temperatura sea al menos aproximadamente 30°C para sondas cortas (por ejemplo, 10 a 50 nucleótidos) y al menos aproximadamente 60°C para sondas largas (por ejemplo, mayores de 50 nucleótidos). Las condiciones de rigurosidad pueden también conseguirse con la adición de agentes desestabilizantes tales como formamida.

El término “anticuerpo” se refiere a un polipéptido codificado sustancialmente por un gen de inmunoglobulina o genes de inmunoglobulina, o fragmentos de los mismos que se unen específicamente y reconocen un analito (antígeno).

Un “anticuerpo quimérico” es una molécula de anticuerpo en la que (a) la región constante, o una parte de la misma, está alterada, remplazada o intercambiada, de modo que el sitio de unión al antígeno (región variable) esté unido a una región constante de una clase, función efectora y/o especie diferente o alterada, o una molécula completamente diferente que otorga nuevas propiedades al anticuerpo quimérico, por ejemplo, una enzima, toxina, hormona, factor de crecimiento, fármaco, etc.; o (b) la región variable, o una parte de la misma, está alterada, reemplaza o intercambiada con una región variable que tiene una especificidad antigénica diferente o alterada.

El término “inmunoensayo” es un ensayo que utiliza un anticuerpo para unirse específicamente a un analito. El inmunoensayo está caracterizado por el uso de propiedades de unión específicas de un anticuerpo particular para aislar, dirigir, y/o cuantificar el analito.

Las expresiones “se une específicamente a una proteína” o “específicamente inmunorreactivo con”, cuando se refiere a un anticuerpo, se refiere a una reacción de unión que es determinante de la presencia de la proteína en presencia de una población heterogénea de proteínas y otros compuestos biológicos. Por tanto, en condiciones de inmunoensayo indicadas, los anticuerpos especificados se unen preferentemente a una proteína particular y no se unen en una cantidad significativa a otras proteínas presentes en la muestra. La unión específica a una proteína en dichas condiciones requiere que un anticuerpo se seleccione por su especificidad por una proteína particular. Puede usarse una diversidad de formatos de inmunoensayo para seleccionar anticuerpos específicamente inmunorreactivos con una proteína particular. Por ejemplo, se usan de forma rutinaria inmunoensayos ELISA en fase sólida para seleccionar anticuerpos monoclonales específicamente inmunorreactivos con una proteína. Véase, Harlow y Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbour Publications, Nueva York, para una descripción de formatos de inmunoensayo y condiciones que pueden usarse para determinar inmunorreactividad específica.

Un “producto génico”, como se usa en este documento, se refiere a un ácido nucleico cuya presencia, ausencia, cantidad o secuencia de ácido nucleico es indicativo de la presencia, ausencia, cantidad, o composición de ácido nucleico del gen. Los productos génicos, por tanto, incluyen, pero sin limitación, un transcrito de ARNm, un ADNc transcrito de forma inversa a partir de un ARNm, y ARN transcrito a partir de ese ADNc, un ADN amplificado a partir del ADNc, un ARN transcrito a partir del ADN amplificado y subsecuencias de cualquiera de estos ácidos nucleicos. Los polipéptidos expresados por el gen o subsecuencias del mismo también son productos génicos. El tipo particular de producto génico será evidente a partir del contexto del uso de los mismos.

Un “fármaco modificado” significa un compuesto que retiene las propiedades farmacéuticas del fármaco original o sustancia activa aunque la estructura del mismo se haya modificado. Además, también se incluye en el término “fármaco” compuestos útiles en métodos de diagnóstico por sus propiedades de unión específica.

Descripción detallada de la invención

En un primer aspecto, la presente invención se refiere a dos proteínas humanas descritas en la Figura 2B, capaces de participar en la vía de PTCH/SHH humana durante el desarrollo embrionario y/o la carcinogénesis, tal como carcinoma de células basales. Las proteínas de acuerdo con la invención están codificadas por variantes de un gen, cuyo ácido nucleico aislado se describe con detalle a continuación y que se denomina *patched* 2 (PTCH2) debido a sus similitudes con *patched* 1 (PTCH1). Por consiguiente, las proteínas de acuerdo con la invención muestran diferencias sustanciales en la secuencia y funciones cuando se comparan con la proteína PTCH1 humana. Las proteínas pueden caracterizarse por sus funciones que cuando se comparan con PTCH1 humana son similares pero distintas de las mismas de ciertas maneras, más específicamente descritas a continuación en la sección “Resultados y análisis”. Las nuevas proteínas humanas PTCH2 de acuerdo con la invención también son distintas de la PTCH2 de ratón aislada previamente. Por tanto, en una realización preferida, comprende las secuencias de aminoácidos descrita en la SEC ID N° 1 y presentada en el Gen-Bank con el número de identificación de proteína AAD17260.1. Una de las presentes proteínas comprende todos los aminoácidos de la secuencia llamada SEC ID N° 1.

Las proteínas de acuerdo con la invención se preparan fácilmente por los especialistas en este campo por técnicas de ADN recombinante usando las moléculas descritas a continuación o cualquier método sintético (véase, por ejemplo, Barany y Merrifield, Solid-Phase Peptide synthesis, pág. 3-284 en The Peptides: Analysis, Synthesis, Biology, Vol. 2: Special Methods in Peptide synthesis, Part A, Merrifield *et al.*, J. Am. Chem. Soc., 2149-2156).

La presente invención también se refiere al uso de los péptidos, polipéptidos y proteínas descritos en este documento como compuestos principales en métodos dirigidos a descubrir nuevas sustancias, es decir, fármacos modificados, tales como sustancias que muestren propiedades equivalentes o incluso más ventajosas que los compuestos principales en sí. Dichos fármacos modificados también pueden diseñarse por métodos de química combinatoria, en la que se diseña específicamente un compuesto estructuralmente similar, por ejemplo, por la ayuda de ordenadores. Como alternativa, el presente fármaco modificado se identifica por exploración de una biblioteca de compuestos candidatos, por ejemplo, usando un anticuerpo de acuerdo con la invención. En el presente contexto, debe entenderse que, cuando se ha identificado dicho fármaco modificado, es posible producirlo por cualquier otra técnica adecuada. Los métodos de proteómica en los que se usan las presentes moléculas, así como también pueden ser útiles para un uso *per se*.

Un segundo aspecto de la invención es un ácido nucleico que codifica una proteína, definida en las reivindicaciones 1 y 2.

La presente solicitud describe el ácido nucleico de PTCH2 genómico humano aislado que comprende partes o toda la secuencia genómica llamada SEC ID N° 5. En la descripción de la secuencia genómica mostrada en la Figura 1, se muestra la estructura de exones/intrones del presente gen. Además de los exones mostrados, también se ha identificado el exón 12a y 12b, definidos específicamente por la SEC ID N° 3 y la SEC ID N° 4, respectivamente. Es interesante observar que hay una variante de corte y empalme que une el exón 12a al segmento 3' del exón 12b conservando los dinucleótidos GT-AG intrónicos. Los exones 12a y 12b no son variantes, pero los exones reales del gen están identificados por secuenciación de la región genómica correspondiente. (Materiales y métodos como se describe a continuación). Por consiguiente, estos descubrimientos muestran que PTCH2 tiene la misma organización estructural de intrones/exones que PTCH1. La presente solicitud describe un transcrito que ha saltado sólo uno de los exones 9 y 10 definidos en la Figura 1. En una realización alternativa, el transcrito de acuerdo con la invención ha saltado tanto el exón 9 como el exón 10. Las variantes de corte y empalme del presente gen se analizan con más detalle a continuación en la sección “Resultados”, todas las cuales se incluyen en el alcance de la presente invención. Este aspecto de la invención posibilita de forma ventajosa diseñar cebadores de PCR adecuados, que a su vez posibilita la exploración de mutaciones de todas las secciones codificantes de los mismos, por ejemplo, por análisis de SSCP, secuenciación, o cualquier otro método adecuado conocido por los especialistas en este campo. Por tanto, el nuevo gen de PTCH2 humano se ha localizado por mapeo de híbridos de radiación en el cromosoma 1p32-35 con D1S211 y WI1404 como los marcadores flanqueantes más cercanos y con una localización estimada 5,5cR de D1S443. Esta región a menudo se pierde por LOH en diversos tipos tumorales diferentes, tales como neuroblastoma, melanoma, cáncer de mama, cáncer de colon, etc. Por consiguiente, PTCH2 es un candidato para un gen supresor de tumores en esta región y la presente invención también abarca métodos de diagnóstico basados en esta nueva descripción.

En esta región cromosómica también se han mapeado tres síndromes de predisposición a cáncer, concretamente, melanoma familiar CMM1, el locus modificador para la poliposis adenomatosa familiar hMom1 y el síndrome de Bebé

Michelin. PTCH2 es también un candidato para el gen que está detrás de estos síndromes hereditarios. Las presentes moléculas, por lo tanto, se usan de forma ventajosa en el contexto de estas afecciones, por ejemplo, en terapia y/o diagnóstico, tal como en ensayos.

Además, la invención también se refiere a diversos cebadores de PCR basados en secuencias sincrónicas, que permiten la amplificación de toda la secuencia codificante. Dichos cebadores se usan de forma ventajosa para la exploración de mutaciones.

La invención también se refiere al uso de los nuevos ácidos nucleicos aislados, por ejemplo, para identificar mutaciones para propósitos de diagnóstico y/o terapéuticos.

Realizaciones adicionales de este aspecto de la invención incluyen sondas de ácido nucleico, por ejemplo, sondas de ADN, ácidos nucleicos marcados, ADNc, ARN, etc., es decir, todos los productos génicos que se pueden obtener por un especialista en este campo en base a las nuevas variantes de PTCH2 humanas aisladas.

Otro aspecto de la invención es un ácido nucleico que corresponde a una cualquiera de las variantes de corte y empalme descritas en la Figura 2B, una proteína o polipéptido codificado por la misma, así como diversos usos de la misma.

Con respecto a la preparación de ácidos nucleicos de acuerdo con la invención, puede usarse cualquier técnica de ADN recombinante o método sintético adecuado. (Para procedimientos de laboratorio generales útiles en este contexto véase, por ejemplo, Sambrook *et al.*, Molecular Cloning, A Laboratory Manual, 2ª ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Berger y Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol 153, Academic Press, Inc., San Diego, CA; Current Protocols in Molecular Biology, F.M. Ausbel *et al.*, eds., Current Protocols (1994)).

Un aspecto adicional de la presente invención es un vector que comprende un ácido nucleico definido anteriormente. Los vectores son, por ejemplo, útiles para transformar células *in vitro* o *in vivo* que expresen las proteínas y péptidos de acuerdo con la invención y pueden ser, por ejemplos, plásmidos, virus, etc.

Otro aspecto de la invención es una célula recombinante, tal como una célula eucariota, por ejemplo, de mamífero, o una célula procariota, por ejemplo, una bacteria, que comprende un vector definido anteriormente. Dichas células, por ejemplo, pueden usarse para controlar los niveles de expresión de las proteínas y polipéptidos de acuerdo con la invención en una amplia diversidad de contextos. Por ejemplo, cuando se tienen que determinar los efectos de un fármaco, se administrará el fármaco al organismo, tejido o célula transformada. Por consiguiente, los sistemas modelo que incluyen dichas células son otro aspecto de la invención.

Un aspecto adicional de la invención es un anticuerpo, tal como un anticuerpo monoclonal o policlonal, que se une específicamente a una proteína o polipéptido de acuerdo con la invención. Una unidad estructural de inmunoglobulina ejemplar (anticuerpo) comprende un tetrámero. Cada tetrámero está compuesto por dos pares idénticos de cadenas polipeptídicas, teniendo cada par una cadena "ligera" (aproximadamente 25 kD) y una cadena "pesada" (aproximadamente 50-70 kD). El extremo N-terminal de cada cadena define una región variable de aproximadamente 100 a 110 o más aminoácidos principalmente responsables del reconocimiento de antígenos. Las expresiones cadena ligera variable (V_L) y cadena pesada variable (V_H) se refieren a estas cadenas ligera y pesada, respectivamente.

La invención también abarca anticuerpos quiméricos u otros anticuerpos que se unen a las presentes proteínas o polipéptidos. Además, la invención también se refiere al uso de los presentes anticuerpos en ensayos. (En este contexto véase, por ejemplo, *Fundamental Immunology*, Third Edition, W.E. Paul, ed., Raven Press, N.Y. 1993).

Además, la invención también se refiere a una célula recombinante que expresa un anticuerpo de acuerdo con la invención.

En general, pueden usarse procariotas para clonar las secuencias de ADN que codifican una cadena de inmunoglobulina humana anti-PTCH2. *E. coli* es un hospedador procariota particularmente útil para clonar las secuencias de ADN de la presente invención. Microbios, tales como levaduras también son útiles para la expresión. *Saccharomyces* es un hospedador de levaduras preferido, teniendo los vectores adecuados secuencias de control de la expresión, un origen de replicación, secuencias de terminación y similares, según se desee. Los promotores típicos incluyen la 3-fosfoglicerato quinasa y otras enzimas glicolíticas. Los promotores de levaduras inducibles incluyen, entre otros, promotores de la alcohol deshidrogenasa 2, isocitocromo C, y enzimas responsables de la utilización de maltosa y galactosa.

Las células de mamífero son un hospedador particularmente preferido para expresar segmentos de nucleótidos que codifican inmunoglobulinas o fragmentos de las mismas (véase, por ejemplo, Winnacker, *From Genes to Clones*, VCH Publishers, N.Y., 1987). Se han desarrollado en la técnica varias líneas celulares hospedadoras adecuadas capaces de secretar proteínas heterólogas intactas, e incluyen las líneas celulares CHO, diversas líneas celulares COS, células HeLa, células L y líneas celulares de mieloma. Preferiblemente, las células son no-humanas. Los vectores de expresión de estas células pueden incluir secuencias de control de la expresión, tales como un origen de replicación, un promotor, un potenciador (Queen *et al.* (1986) *Immunol. Rev.* 89:49), y sitios con información de procesamiento necesarias,

tales como sitios de unión al ribosoma, sitios de corte y empalme del ARN, sitios de poliadenilación, y secuencias terminadoras de la transcripción. Las secuencias de control de la expresión preferidas son promotores derivados de genes endógenos, citomegalovirus, SV40, adenovirus, papilomavirus bovinos y similares (véase, por ejemplo Co *et al.* (1992) *J. Immunol.* 1458:1149).

Un aspecto adicional de la presente invención es un kit para la detección de un gen o polipéptido de PTCH2 humano que comprende en un recipiente una molécula seleccionada entre el grupo compuesto por un ácido nucleico, un polipéptido o una proteína o un anticuerpo de acuerdo con la invención. Los componentes adecuados adicionales de dicho kit se determinan fácilmente por los especialistas en este campo así como las condiciones para el uso del mismo.

Además, la invención también se refiere al uso de un ácido nucleico descrito en la reivindicación 1 ó 25 en terapia génica. Además de dichas secuencias descritas específicamente, uno cualquiera de los exones descritos en este documento puede usarse para este fin. Para una revisión de los procedimientos de terapia génica véase Anderson, *Science* (1992) 256:808-813; Nabel y Felgner (1993) *TIBTECH* 11:211-217; Mitani y Caskey (1993) *TIBTECH* 11:162-166; Mulligan (1993) *Science* 926-932; Dillon (1993) *TIBTECH* 11: 167-175; Miller (1992) *Nature* 357:455-460; Van Brunt (1988) *Biotechnology* 6(10): 1149-1154; Vigne (1995) *Restorative Neurology and Neuroscience* 8: 35-36; Kremer y Perricaudet (1995) *British Medical Bulletin* 51(1) 31-44; Haddada *et al.* (1995) en *Current Topics in Microbiology and Immunology* Doerfler y Böhm (eds) Springer-Verlag, Heidelberg Alemania; y Yu *et al Gene Therapy* (1994) 1:13-26.

El suministro del gen o el material génico al interior de la célula es la primera etapa crítica en el tratamiento con terapia génica de una enfermedad. Los especialistas en la técnica conocen bien una amplia diversidad de métodos de suministro. Dichos métodos incluyen, por ejemplo, suministro de genes basado en liposomas (Debs y Zhu (1993) documento WO 93/24640; Mannino y Gould-Fogerite (1988) *Bio Techniques* 6(7):682-691; Rose Patente de Estados Unidos N° 5.279.833; Brignam (1991) documento WO 91/06309; y Felgner *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84: 7413-7414), y vectores retrovirales deficientes en replicación que tienen una secuencia polinucleotídica terapéutica como parte del gen retroviral (véase, por ejemplo, Miller *et al.* (1990) *Mol. Cell. Biol.* 10:4239 (1990; Kolberg (1992) *J. NIH Res.* 4:43, y Cornetta *et al. Hum. Gene Ther.* 2:215 (1991)). Los vectores retrovirales ampliamente usados incluyen los basados en el virus de leucemia murina. (MuLV), virus de la leucemia del mono gibbon (GaLV), virus de la inmunodeficiencia del simio (SIV), virus de la inmunodeficiencia humana (VIH), y combinaciones de los mismos. Véase, por ejemplo, Buchscher *et al.* (1992) *J. Virol.* 66 (5) 2731-2739; Johann *et al* (1992) *J. Virol.* 66 (5):1635-1640 (1992); Sommerfelt *et al.*, (1990) *Virol.* 176:58-59; Wilson *et al* (1989) *J. Virol.* 63:2374-2378; Miller *et al.*, *J. Virol.* 65:2220-2224 (1991); Wong-Staal *et al.*, documento PCT/US94/05700, y Rosenberg y Fauci (1993) en *Fundamental Immunology, Third Edition* Paul (ed) Raven Press, Ltd., Nueva York y las referencias del mismo, y Yu *et al.*, *Gene Therapy* (1994) *supra*.

La presente invención también puede usarse en la industria farmacéutica. Por ejemplo, se proporcionará información que pueda posibilitar finalmente que las células del tejido fetal puedan transplantarse en pacientes que padecen, por ejemplo, enfermedad de Parkinson o cáncer, tal como BCC. (Para una breve revisión de los métodos de suministro de fármacos véase Langer 249:1 527-1533 (1990), Remington's Pharmaceutical Sciences, Mack Publishing Company, Philadelphia, PA, 17ª ed. (1985) etc.)

Descripción detallada de los dibujos

La Figura 1 muestra la secuencia genómica de la SEC ID N° 5, en la que los exones y los intrones están indicados en la secuencia genómica del presente gen patched 2 humano. Sin embargo, los exones 12a y 12b analizados anteriormente no se muestran específicamente en la Figura 1, pero en su lugar se describen como secuencias diferentes SEC ID N° 3 y SEC ID N° 4, respectivamente. La Figura 2A describe una comparación de secuencia de aminoácido de la secuencia de PTCH2 (líneas superiores) y PTCH1 (líneas inferiores) humanas. Las líneas verticales indican aminoácidos idénticos, mientras que los puntos aminoácidos similares. La secuencia de PTCH2 presentada está compuesta por los clones de ADNc originales y los productos del análisis RACE 5'.

La Figura 2B es una representación de los acontecimientos de corte y empalme alternativo que producen diferentes extremos C-terminales. En la glándula parótida y el colon, el penúltimo exón y el último exón están canónicamente unidos entre sí. En cerebro fetal, sin embargo, el penúltimo exón con parte del intrón 3' funciona como el exón terminal. La secuencia intrónica se muestra por letras minúsculas y la exónica flanqueante por letras mayúsculas. Por encima de la secuencia de nucleótidos, se muestra la secuencia de aminoácidos deducida, y por debajo se encuentra la secuencia correspondiente de Ptc2 de ratón. Los dinucleótidos intrónicos conservados se muestran por letras en negrita y las señales de terminación se indican por asteriscos. Obsérvese la ausencia de conservación de la posición de los codones de terminación entre la secuencias de PTCH2 de ratón y humana. Las supuestas señales de poliadenilación también se muestran en este diagrama. La organización genómica se obtuvo analizando clones BAC que incluyen el gen de PTCH2.

La Figura 2C es una representación de las diferentes variaciones de transcritos con corte y empalme que incluyen las secuencias del exón 1 y el exón 2. Los exones 1 y 2 canónicos se muestran por recuadros y el intrón entre ellos por una línea continua. Los dinucleótidos GT y AG que abarcan las secuencias que se usan como intrones en transcritos individuales se indican por letras minúsculas. G, estructura genómica, obtenida de los segmentos de secuenciación de

clones BAC que incluyen el gen de PTCH2; C, transcrito canónico; A, transcrito A (los exones 9 y 10 saltados de este producto no se muestran en el diagrama); B, transcrito B.

La Figura 3A es una fotomicrografía de campo oscuro del tumor BCC hibridado con la sonda antisentido marcada con ³⁵S que muestra la señal abundante para el ARNm de PTCH1 (granos brillantes) en todas las células tumorales BCC.

La Figura 3B describe la sobreexpresión del ARNm de PTCH2 en BCC y se expresa principalmente en contraste en las células basaloides en la periferia de las redes tumorales.

La Figura 3C es otro BCC que muestra una fuerte señal de ARNm de PTCH2 en la periferia de la red tumoral (Tu), mientras que no se detecta señal en la epidermis (Ep).

La Figura 3D son secciones del mismo tumor (C) hibridado con la sonda con sentido de PTCH2 que no muestra señal.

La Figura 3E muestra inmunorreactividad para Ki-67 (precipitado marrón) observado en la periferia, en las células que mostraron regulación positiva fuerte del ARNm de PTCH2.

La Figura 3F describe redes tumorales bajo un aumento de alta potencia que muestra una abundante señal de ARNm de PATCH2 (granos negros) en las células tumorales basaloides oscuras y una señal inferior en el centro (flecha). Barras (A-E), 24 µm, y F, 6 µm.

Parte experimental

Materiales y métodos

En el presente contexto, se hace una referencia general a G. Zaphiropoulos *et al.*, Cancer Res., vol. 59, pág. 787-792, 15 de febrero de 1999, que describe métodos útiles en el presente contexto. Todas las referencias mencionadas en la presente solicitud se incluyen por la presente en este documento como referencia. Los siguientes ejemplos no pretenden limitar el alcance de la invención sino que son simplemente una ilustración.

Los análisis RACE se realizaron esencialmente como se ha descrito anteriormente (Zaphiropoulos, P.G. y Toftgard, R. (1996): "cDNA cloning of a novel WD repeat protein mapping to the 9q22.3 chromosomal region", DNA Cell Biol. **15**, 1049-1056) usando el kit Marathon (Promega). Las secuencias cebadoras usadas para RACE están disponibles a petición.

Las sondas de ARN marcadas con ³⁵S de PTCH2 usadas para las hibridaciones *in situ*, que se realizaron como se ha descrito previamente (Unden *et al.*, (1997) *supra*), corresponden a las posiciones 218 a 437 y 838 a 920 de la secuencia de PTCH2 de la SEC ID N° 1.

Resultados y análisis

Para identificar componentes adicionales de la cascada de PTCH/SHH de acontecimientos de señalización, se buscó en la base de datos Incyte LifeSeq™ (Incyte Pharmaceuticals Inc., Palo Alto, CA, USA) usando las secuencias de PTCH. Además de los clones que representan el ADNc de PTCH, se identificaron dos ADNc casi idénticos, de la glándula parótida y el colon, que contenían secuencias similares a, pero distintas de, el extremo 3' de PTCH. Por análisis RACE 5' usando ADNc de cerebro fetal se obtuvo información de secuencia adicional de estos transcritos (llamados PTCH2) y que corresponde al ADNc de longitud completa (Figura 2A). PTCH2 es idéntica al 57% con PTCH1, con una región significativamente variable presente entre los dominios transmembrana 6 y 7, e idéntica al 91% a la secuencia de Ptc2 de ratón recientemente publicada (Motoyama, J., Takabatake, T., Takeshima, K. y Hui, C. (1998): "Ptc2, a second mouse Patched gene is co-expressed with Sonic hedgehog", Nature Genet, **18**, 104-106). Similar al gen de ratón, PTCH2 carece de la extensión C-terminal presente en la PTCH1 de seres humanos, de ratón y de pollo (Goodrich, L.V., Johnson, R.L., Milenkovic, L., McMahon, J.A. y Scott, M.P. (1996): "Conservation of the hedgehog/patched signalling pathway from flies to mice: Induction of a mouse patched by Hedgehog", Genes Dev. **10**, 301-312, Mariago, V., Scott, M.P., Johnson, R.L., Goodrich, L.V. y Tabin, C.J. (1996): "Conservation in hedgehog signalling: Induction of a chicken patched homolog by Sonic hedgehog in the developing limb" Development **122**, 1225-1233). Sin embargo, de acuerdo con la presente invención, se ha demostrado que el ADNc de PTCH2 humana termina 36 aminoácidos antes que la secuencia de Ptc2 de ratón. Además, cuando se realizó RACE 3' a partir del cerebro fetal, se identificó una región C-terminal alternativa. Esto tuvo una gran similitud estructural con la secuencia C-terminal de Ptc2 de ratón y se origina de la región genómica que se une a al menos 2 exones de PTCH2 (Figura 2B). Por lo tanto, en estos transcritos con corte y empalme alternativo, el penúltimo exón con un segmento del intrón 3' contiguo sirve como exón terminal.

Además, los transcritos humanos y de ratón diferían en la posición de las señales de terminación (la secuencia humana es 21 aminoácidos más larga), lo que sugiere una función no conservada, específica de la especie de este dominio C-terminal alternativo. El descubrimiento de dos posibles regiones C-terminales para PTCH2 es intrigante e implica un papel de este fenómeno en la modulación de la señalización. Los transcritos con corte y empalme alternativo

adicionales también se identificaron por el análisis RACE (Figura 2C). El transcrito A carece de la secuencia que corresponde a los exones 9 y 10 de PTCH1 (comparaciones preliminares de las uniones intrón-exón de PTCH2 con PTCH1 indican una organización genómica similar), estando la fase de lectura abierta retenida en la unión del exón 8 al exón 11. Los exones 9 y 10 codifican la última parte del primer bucle extracelular y los dominios transmembrana 2 y 3 en la supuesta estructura de la proteína PTCH1. Además, este transcrito también carece de un segmento 5' del exón canónico 2, debido al uso de un sitio de corte y empalme 3' alternativo presente en este exón, manteniéndose la fase de lectura abierta. La consecuencia funcional de este corte y empalme alternativo aún no se conoce, pero es interesante observar que los bucles extracelulares en PTCH1 se supone que están implicados en la unión del ligando SHH (Marigo *et al.*, (1996), *Nature* **384**, *supra*; Stone *et al.*, (1996), *Nature* **384**, *supra*) y que la inserción de un nuevo casete en el intrón 9 del gen de PTCH1 de ratón está asociado con un fenotipo grave (Hahn, H., Wojnowski, L., Zimmer, A.M., Hall, J., Miller, G. y Zimmer, A. (1998): "Rhabdomyosarcomas and radiation hypersensitivity in a mouse model of Gorlin syndrome", *Nature Med.* **4**, 619-622). Además, los exones 9 y 10 codifican parte del supuesto dominio sensible a esteroides (Osborne, T.F. y Rosenfeld, J.M. (1998): "Related membrane domains in proteins of sterol sensing and cell signalling provide a glimpse of treasures still buried within the dynamic realm of intracellular metabolic regulation", *Curr. Opin. Lipidol.* **9**, 137-140) también encontrado en PTCH1, y que se ha implicado recientemente en la mediación del efecto modulador potente del colesterol sobre la señalización de SHH/PTCH (Cooper, M.K., Porter, J.A., Young, K.E., y Beachy, P.A. (1998): "Teratogen-mediated inhibition of target tissue response to Shh signalling" *Science* **280**, 1603-1607). Por tanto, si PTCH2 también sirve como receptor para SHH y/o factores relacionados, la forma del receptor que carece de los exones 9 y 10 puede mostrar propiedades de señalización alteradas. El transcrito B contiene secuencias adicionales entre los exones canónicos 1 y 2, que se originan a partir del extremo 5' del intrón 1. La fase de lectura abierta que incluye la metionina iniciadora del exón 1 no se mantiene en este transcrito, lo que sugiere que, si este transcrito es funcional, la metionina del exón 2 o los codones que no son metionina se usan para producir un producto proteico, en similitud con lo que se ha propuesto para los productos con corte y empalme alternativo de PTCH1 humana (Hahn *et al.*, *Cell* **85**, *supra*). Por mapeo de híbridos de radiación, se localizó el gen de PTCH2 en el brazo corto del cromosoma 1, a diferencia de PTCH1 que reside en el cromosoma 9q22.3.

Se ha informado de que los homólogos de ratón y pez cebra de PTCH2 se expresan en un patrón parcialmente solapante con PTCH1 durante el desarrollo embrionario y que están inducidos por SHH (Motoyama *et al.*, (1998) *Nature Genet.* **18**, *supra*, Concordet, J.P., Lewis, K.E., Moore, J.W., Goodrich, L.V., Johnson, R. L., Scott, M.P., y Ingham, P.W. (1996): "Spatial regulation of a zebrafish patched homologue reflects the roles of sonic hedgehog and protein kinase A in a neural tube and somite patterning", *Development* **122**, 2835-2846), lo que implica un papel en esta vía de señalización. Con estos antecedentes fue interesante analizar la expresión de PTCH2 en BCC que muestran regulación positiva uniforme de PTCH1 en todas las células tumorales (Undén *et al.*, (1997) *Cancer res.* **57**, *supra*). Se realizó hibridación *in situ* sobre seis BCC familiares y cuatro BCC esporádicos de diferentes subtipos histológicos. Se observó una fuerte señal positiva para el ARNm de PTCH2 exclusivamente en las células tumorales de todos los BCC. De forma notable, la señal fue uniformemente más fuerte en las células periféricas en empalizada de las redes tumorales (Figura 2). Estas células también mostraron una inmunotinción positiva para el marcador de proliferación celular Ki-67.

El descubrimiento de que en BCC que tienen mutaciones frecuentes en el gen de PTCH1, la expresión de los ARNm de PTCH2 está regulada positivamente, une ligeramente la nueva PTCH2 de acuerdo con la invención con la cascada de PTCH/SHH de acontecimiento de señalización. Por lo tanto, es probable que PTCH2 represente un gen diana de esta vía que está bajo la regulación negativa de PTCH1, precisamente como PTCH1 en sí misma. Además, esta observación sugiere fuertemente que PTCH2 tiene funciones diferentes a PTCH1 ya que la regulación positiva de la expresión de PTCH2 parece incapaz de compensar la proteína PTCH1 inactiva. Esta conclusión también está apoyada por la letalidad embrionaria temprana observada en ratones PTCH1 (-/-) 5,13 y la ausencia de heterogenicidad genética en el síndrome de Gorlin. Sin embargo, si PTCH2 puede bloquear la señalización constitutiva de SMO, o pudiera funcionar como receptor de SHH adicional, dependiendo posiblemente del corte y empalme alternativo, permanece como objeto de experimentación adicional.

REIVINDICACIONES

1. Una secuencia de ácido nucleico aislada que codifica una proteína Ptch2 humana que tiene la SEC ID N° 1.

2. Una secuencia de ácido nucleico aislada que codifica una variante de corte y empalme que tiene los aminoácidos 1-1142 de la SEC ID N° 1 inmediatamente seguida por la secuencia TrpGlyAlaSerSerSerLeuProGlnSerPheAlaArgValThrThrSerMetThrValAlaIleHisProProProLeuProGlyAlaTyrIleHisProAlaProAspGluProProTrpSerProAlaAlaThrSerSerGlyAsnLeuSerSerArgGlyProGlyProAlaThrGly.

3. Una proteína que tiene la SEC ID N° 1 o que tiene los aminoácidos 1-1142 de la SEC ID N° 1 inmediatamente seguida por la secuencia TrpGlyAlaSerSerSerLeuProGlnSerPheAlaArgValThrThrSerMetThrValAlaIleHisProProProLeuProGlyAlaTyrIleHisProAlaProAspGluProProTrpSerProAlaAlaThr SerSerGlyAsnLeuSerSerArgGlyProGlyProAlaThrGly.

4. Una secuencia de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1-2, o una proteína de acuerdo con la reivindicación 3, para su uso como medicamento.

5. Uso de una secuencia de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1-2, o una proteína de acuerdo con la reivindicación 3, para la fabricación de un medicamento para el tratamiento de una afección que implique tumores, tal como BCC.

6. Un método de diagnóstico *in vitro*, en el que se usa una secuencia de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1-2, o una proteína de acuerdo con la reivindicación 3.

7. Un método para explorar, en el que se explora una biblioteca de compuestos candidatos adecuados para fármacos modificados, usando una proteína de acuerdo con la reivindicación 3 como compuesto principal.

8. Un vector que comprende una secuencia de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1-2.

9. Una célula recombinante que comprende un vector de acuerdo con la reivindicación 8.

10. Un anticuerpo que se une específicamente a una proteína de acuerdo con la reivindicación 3.

11. Una célula recombinante que expresa un anticuerpo de acuerdo con la reivindicación 10.

12. Un kit para la detección de un gen o proteína de Ptch2 humana, que comprende en un recipiente una molécula seleccionada entre el grupo compuesto por una secuencia de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1-2, una proteína de acuerdo con la reivindicación 3, o un anticuerpo de acuerdo con la reivindicación 11.

13. Uso de una secuencia de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1-2, para la fabricación de un medicamento para su uso en terapia génica.

14. Uso *in vitro* de una secuencia de ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 1-2, como sonda, un cebador o un reactivo de diagnóstico.

Faltan las secuencias intrónicas entre los exones 2 - 3 y los exones 18 - 19 (intrones: letras minúsculas, exones: letras mayúsculas). Las letras minúsculas en el primer exón indican los nucleótidos que no se han determinado de forma inequívoca.

Exón 1

```

1  CGGGTGAATC CCGGCGCCGC GCCCGGACC CGCAGCTCCC TGCCTCTCTC
51  CCTCCCAGCC GCTTTAACAC CCACACCCCA CAGTCTCTCC CACGsCCGCG
101 CCTTGGCGGC CCCACTGAAT CCCTACGCGG GGCCAGCGG TACCGGGAGA
151 CCGGGCTAGC CTATGGGAGC GCCCAGATAA CGCGGGTTGG GGGCGCCCGC
201 GCGGcCATCC CCGCCAGCAT GACTCGATCG CCGCCCCTCA GAGAGCTGCC
251 CCCGAGTTAC ACACCCCCAG CTCGAACCGC AGCACCCCAG gtgagtagag
301 ggggagctgg aagaaggaag agagcggagc caggtctgtc actcgggcct
351 ctgcaagggtt tgtgatgtct tgaagtgccg agtgtcatta gatgtctgaa
401 ggcaagttag agccagcacc gcaagcaagt tgtgcgtgtg tgtcgggtgtg
451 tctgtgccgg tgtctcctca tcgtctggcc agtgagaatg aatgtctgtg
501 ggttcacctc tgtgtccacc cgacgacagg tgtgtgtaca tatgtatcct
551 gctctcagaa aatgggccta tgccgocggg cgcggtgact cacgcctgta
601 atcccaacac tgggaggctg aggcaggcag attacctgag gtcaggagtt
651 cgagaccagc caggccaaca tggggaaact ctgtctctac taaaaataaa
701 aattagcagg gcgtggtggc gggcgccctgt agtccaaact actcgggagg
751 ctgaggcagg agaattctctt gaacctggga ggcgagggtt gcagtcaagc
801 cgagatcaca ccactgcact ccagccaggg caacagagcg agatgcgtct
851 caaaaaaaaa aaaaaaaaaa aaaaggagag aaaacaaaaa gaaaagaaag
901 gaaaataggc ctatgccttc ctcagggtgtg tgctggggat ggtgggtgtt
951 acatcttcca agtctgggcc tgtgtctgtg ttggtgctcc ctgtcccaca
1001 tccagaaatc aagaagcagag ggctgggcag cagatatata gggtgagaag

```

Fig. 1

1051 ggaaggattt catgcattgt tacagtgatg cctggctgac ccttctcttt
 EXON 2
 1101 ccatcccagA TCCTAGCTGG GAGCCTGAAG GCTCCACTCT GGCTTCGTGC
 1151 TTA~~CT~~TCCAG GGCCTGCTCT TCTCTCTGGG ATGCGGGATC CAGAGACATT
 1201 GTGGCAAAGT GCTCTTTCTG GGACTGTTGG CCTTTGGGGC CCTGGCATTG
 1251 GGTCTCCGCA TGGCCATTAT TGAGACAAAC TTGGAACAGC TCTGGGTAGA
 1301 AGTGGGCAGC CGGGTGAGCC AGGAGCTGCA TTACACCAAG GAGAAGCTGG
 1351 GGGAGGAGGC TGCATACACC TCTCAGATGC TGATACAGAC CGCACGCCAG
 1401 GAGGGAGAGA ACATCCTCAC ACCCGAAGCA CTTGGCCTCC ACCTCCAGGC
 1451 AGCCCTCACT GCCAGTAAAG TCCAAGTATC ACTCTATGGG AAG.....
 1501
 1551 tgagtctggc tgagcccctg agcagctggg ggcgaggcgt gctgtggggg
 1601 ttctggagtg ggaatcccct tcttctgctg atctcctatg cccctggcta
 EXON 4
 1651 ttgcagTCCT GGGATTGAA CAAAATCTGC TACAAGTCAG GAGTTCCCCT
 1701 TATTGAAAAT GGAATGATTG AGCGGgtaag tgtcctgaga gggagtagag
 1751 gcagaacttt ttctgtagcg tgggaggact cagagaccga gcaagcccca
 1801 cagcctgcaa tctgccccct taaaactaag gagggggatt gcagagggca
 1851 tcctacaaag gttgtggggc aggactgacg tggcccgggg tatccctggc
 EXON 5
 1901 agATGATTGA GAAGCTGTTT CCGTGCGTGA TCCTCACCCC CCTCGACTGC
 1951 TTCTGGGAGG GAGCCAAACT CCAAGGGGGC TCCGCCTACC TGCCgtgagt
 2001 gccactcctg gggccctgct tcatctcccg ctggggactc tcccagcaga
 2051 aaggaggggt ctggggaatg aggatgatca aaaccttacc aaggtcctaa
 2101 ttacctccca ggccaggaac agagagcatg ggcttcccca aggctctctc
 2151 cacatcctcc ttctctttcc ctctcaagga aggaagacct gacttattta
 2201 cacaaaacta aacacaaaga tctgtaagat ctgagcaaag gagaaaaaga
 2251 tccccacaaa gaggctttgc tgggggaaat tacctaggtg tttgctaagc
 2301 cattgcccag gccagaaaga aaacctgcta caggcatgtg cctgctggtt
 2351 gtatattaga accaagcaca cagcttggtg aggaactcag tggggccttt

Fig. 1 (cont.)

2401 ctggggccctt tctatgtatt aggtaaccct gccctgatat tegtctcagc
 2451 cccttgtact cttctacagc tcaactgtagc accctgggtg gcccatgcag
 2501 cctggcagtt ctgagaagct gaggcttgca caccctccat atggaaggac
 2551 aaatcggcag ataagaggag ggtgggggtac agcatggcgc cccagcagca
 2601 gtttggagcc tgggttttcg tccctgaccc tcaccaacta taggcttttc
 2651 cctcagCGGC CGCCCGGATA TCCAGTGGAC CAACCTGGAT CCAGAGCAGC
 2701 TGCTGGAGGA GCTGGGTCCC TTTGCCTCCC TTGAGGGCTT CCGGGAGCTG
 2751 CTAGACAAGG CACAGGTGGG CCAGGCCTAC GTGGGGCGGC CCTGTCTGCA
 2801 CCCTGATGAC CTCCACTGCC CACCTAGTGC CCCCACCAT CACAGCAGGC
 2851 AGgtgggttc caaccaggtc tgccagggaa aggctgtttt ccttcccttt
 2901 cccttctctca tactctctgtg ttctggggga gctgactgct ctgtgccttg
 2951 accccccact tcttgcccat tattaccctg ctcccacagt gccaggcccc
 3001 caatgttcca ttcccattca gttatcctac ggagccctca agtggatatat
 3051 atgaatccct ttttctttt ctaagcctag ataaggctgg acttcttttt
 3101 tttttttttt ttgagtctca ctctgtcacc caggctggag tgcagtagtt
 3151 cgatcttgge tcaactgcaac ctcggtcaa gcaattctcc tgccttagcc
 3201 tcttgagtag ctgggattac aggtgccac caccatgccc ggctaatttt
 3251 tattagcctc ccaaagtgtc gggattacag gcgtgagcca ctgcgcctgg
 3301 ccaaggctgg actttttatc aaaatagact aatacaggga aactaagaac
 3351 acagcaggta agcatgaata tcatacctgg tttcccagggt ttctttgtgg
 3401 ccctgcaa at gtggtacttt tttcagaatc cgccagttac accagctcct
 3451 ccagaagcc tacttccagg cctctgtctc cccttggggc ttctgtctg
 3501 cgggatacta gctgttcaact cctgcagagc agtcaagagg ctccagaatag
 3551 ttacctacac tccagcccta ctgagcttca tggcagcgtg gttcctggag
 3601 gtggaagccc agggacactc agttatccac ggccagggcc ttgagcatta
 3651 accctcctg ttccctctca gGGCTCCCAA TGTGGCTCAC GAGCTGAGTG
 3701 GGGGCTGCCA TGGCTTCTCC CACAAATTCA TGCAC TGGCA GGAGGAATTG

Fig. 1 (cont.)

3751 CTGCTGGGAG GCATGGCCAG AGACCCCCAA GGAGAGCTGC TGAGgtaggg
 3801 tctcctctgg gagttggtga ggggactctg ttcattgagaa cccatactgt
 3851 aatgccaggc agctctggca aaaggccctt cacatccctc accaggtggt
 3901 tgggccagct ctgacccttg gttctccac acccccacca gGGCAGAGGC
 3951 CCTGCAGAGC ACCTTCTTGC TGATGAGTCC CCGCCAGCTG TACGAGCATT
 4001 TCCGGGGTGA CTATCAGACA CATGACATTG GCTGGAGTGA GGAGCAGGCC
 4051 AGCACAGTGC TACAAGCCTG GCAGCGGCGC TTTGTGCAGg tcggtatgga
 4101 caaggacaag gggggtgccc tgaggccatt cctcctcct gcccctcct
 4151 atccaccctg tttctccagc TGGCCAGGA GGCCCTGCCT GAGAACGCTT
 4201 CCCAGCAGAT CCATGCCCTC TCCTCCACCA CCCTGGATGA CATCCTGCAT
 4251 GCGTTCTCTG AAGTCAGTGC TGCCCGTGTG GTGGGAGGCT ATCTGCTCAT
 4301 Ggtgggtctt gcacctgga ccttgcccc accccacctc caaccagtgc
 4351 ccacctggg agccctgag actgccctt cccccacag CTGGCCTATG
 4401 CCTGTGTGAC CATGCTGCGG TGGGACTGCG CCCAGTCCCA GGGTTCCGTG
 4451 GGCCTTGCCG GGGTACTGCT GGTGGCCCTG GCGGTGGCCT CAGGCCTTGG
 4501 GCTCTGTGCC CTGCTCGGCA TCACCTTCAA TGCTGCCACT ACCCAGgtac
 4551 gccaggactg cagggcagac tcagtgccag tcaccaggct tcacgggtcc
 4601 tcagctgccc gctcctctgc cctccagGT GCTGCCCTTC TTGGCTCTGG
 4651 GAATCGGCGT GGATGACGTA TTCCTGCTGG CGCATGCCCTT CACAGAGGCT
 4701 CTGCCCTGGCA CCCCTCTCCA Ggtggggcct tgtccccag ggctcatctg
 4751 aggcagctca gcttactggt taagagcctc ttggttcaag tgacccttg
 4801 gctgctaatag aacctcggtg cctcttgctc ccatctgtaa acaggggaaa
 4851 taatagtgt gtgtcctaag ggttattgtt tggatcagtg aggtaactca
 4901 agttgaatgc ttagaacagc ccatcatacg tacatggtac ccaataaatg
 4951 ctaggcactg tggtatgact gcccacctc tgcaccccaa gttcctgagc
 5001 ctccccttca ctccactttg acacggcccc tcccttgatga cctgagggca
 5051 ggtccccact ctgtcctggc agGAGCGCAT GGGCGAGTGT CTGCAGCGCA

Fig. 1 (cont.)

5101 CGGGCACCAG TGTCGTACTC ACATCCATCA ACAACATGGC CGCCTTCCTC
 5151 ATGGCTGCCC TCGTTCCCAT CCCTGCGCTG CGAGCCTTCT CCCTACAGGC
 5201 GGCCATAGTG GTTGGCTGCA CCTTTGTAGC CGTGATGCTT GTCTTCCCAG
 5251 GCATCCTCAG CCTGGACCTA CGGCGGCGCC ACTGCCAGCG CCTTGATGTG
 5301 CTCTGCTGCT TCTCCAGgta ctgcgtgcgc cccagcccct tcctcccgtg
 5351 acccacgcca gcctgtcccc tcaccagcat ttcaaggcac agacctgtca
 5401 tccactctct acctcttcca gTCCCTGCTC TGCTCAGGTG ATTCAGATCC
 5451 TGCCCCAGGA GCTGGGGGAC GGGACAGTAC CAGTGGGCAT TGCCCCACCTC
 5501 ACTGCCACAG TTCAAGCCTT TACCCACTGT GAAGCCAGCA GCCAGCATGT
 5551 GGTCACCATC CTGCCTCCCC AAGCCCACCT GGTGCCCCCA CCTTCTGACC
 5601 CACTGGGCTC TGAGCTCTTC AGCCCTGGAG GGTCCACACG GGACCTTCTA
 5651 GGCCAGGAGG AGGAGACAAG GCAGAAGGCA GCCTGCAAGT CCCTGCCCTG
 5701 TGCCCGCTGG AATCTTGCCC ATTTGCGCCG CTATCAGTTT GCCCCGTTGC
 5751 TGCTCCAGTC ACATGCTAAG gtaagactgg gcagagcagg gcagagactt
 5801 agcatctctg ggcccagaag ggcagagagg gcttagtcca ctgcctgagg
 5851 ggctgggggc agccctgggg tctccagctt agttgctaca tcccgcagGC
 5901 CATCGTGCTG GTGCTCTTTG GTGCTCTTCT GGGCCTGAGC CTCTACGGAG
 5951 CCACCTTGGT GCAAGACGGC CTGGCCCTGA CGGATGTGGT GCCTCGGGGC
 6001 ACCAAGGAGC ATGCCTTCCT GAGCGCCCAG CTCAGGTACT TCTCCCTGTA
 6051 CGAGGTGGCC CTGGTGACCC AGGGTGGCTT TGACTACGCC CACTCCCAAC
 6101 GCGCCCTCTT TGATCTGCAC CAGCGCTTCA GTTCCCTCAA GGCGGTGCTG
 6151 CCCCCACCGG CCACCCAGGC ACCCCGCACC TGGCTGCACT ATTACCGCAA
 6201 CTGGCTACAG Ggtgagaggg gaggagacgg gcagggaggg gtgctgcagg
 6251 gagaaacgcc ctggggccac cagctaataag aaccctatcc tggctctccc
 6301 cagGAATCCA GGCTGCCTTT GACCAGGACT GGGCTTCTGG GCGCATCACC
 6351 CGCCACTCGA CCGCAATGGC TCTGAGGATG GGGCCCTGGC CTACAAGCTG
 6401 CTCATCCAGA CTGGAGACGC CCAGGAGCTT CTGGATTTCa GCCAGgttgg

Fig. 1 (cont.)

6451	gagaggggctg	gagggggtcca	ctagtagacagg	ggctgcaggc	ctcctgggcc
6501	caggccttca	gccctctctg	cctctgcagc	EXÓN 16 TGACCACAAG	GAAGCTGGTG
6551	GACAGAGAGG	GACTGATTCC	ACCCGAGCTC	TTCTACATGG	GGCTGACCGT
6601	GTGGGTGAGC	AGTGACCCCC	TGGGTCTGGC	AGCCTCACAG	GCCAACTTCT
6651	ACCCCCCACC	TCCTGAATGG	CTGCACGACA	AATACGACAC	CACGGGGGAG
6701	AACTTTCGCA	gtgagtcttg	gggggagctc	ggcaagagcc	tcagcctcgc
6751	ccacacaagc	cctgagcctg	aggeccctgcc	cactctgccc	cgtgctcacc
6801	gccctgtccc	tctccctctt	ctcccttccc	ctccccctcca	EXÓN 17 cagTCCCCGC
6851	AGCTCAGCCC	TTGGAGTTTG	CCCAGTTCCC	TTTCCTGCTG	CGTGGCCTCC
6901	AGAAGACTGC	AGACTTTGTG	GAGGCCATCG	AGGGGGCCCCG	GGCAGCATGC
6951	GCAGAGGCCG	GCCAGGCTGG	GGTGCACGCC	TACCCCAGCG	GCTCCCCCTT
7001	CCTCTTCTGG	GAACAGTATC	TGGGCCTGCG	GCGCTGCTTC	CTGCTGGCCG
7051	TCTGCATCCT	GCTGGTGTGC	ACTTTCCTCG	TCTGTGCTCT	GCTGCTCCTC
7101	AACCCCTGGA	CGGCTGGCCT	CATAgtgagt	gcttgcagga	gtggggacag
7151	agacacccca	cccttccctg	cccagcctgt	catccctcct	gccaggagcc
7201	ctctgtgagc	cctgtctccc	EXÓN 18 tcagGTGCTG	GTCCCTGGCGA	TGATGACAGT
7251	GGAACTCTTT	GGTATCATGG	GTTTCCTGGG	CATCAAGCTG	AGTGCCATCC
7301	CCGTGGTGAT	CCTTGTGGCC	TCTGTAGGCA	TTGGCGTTGA	GTTACAGTC
7351	CACGTGGCTC	TGGGCTTCCT	GACCACCCAG	GGCAGCCGGA	ACCTGCGGGC
7401	CGCCCATGCC	CTTGAGCACA	CATTTGCCCC	CGTGACCGAT	GGGGCCATCT
7451	CCACATTGCT	GGGTCTGCTC	ATGCTTGCTG	GTTCCCACTT	TGACTTCATT
7501	GTAAG.....				
7551		gtagggaggg	ctcggggcag	ggaggcaggg	ctcaggacag
7601	gcctgggctg	actcccccca	caccctaccc	EXÓN 20 ctagGTACTT	CTTTGCGGCG
7651	CTGACAGTGC	TCACGCTCCT	GGGCCTCCTC	CATGGACTCG	TGCTGCTGCC
7701	TGTGCTGCTG	TCCATCCTGG	GCCCCCGGCC	AGAGgtgacc	acaccctcgg
7751	caccatecct	ctaactcccag	cccaagggac	ggggtaggga	gaggcaaggg

Fig. 1 (cont.)

```

7801 aagggacaga gccctgtggc ccacagacag gtacctcccc aacaggtgcc
7851 accagctgaa ggtggcagcc tcctcctttc cccagacacc atgttcctgc
7901 ccctcagccc tcctggcttc ttcattgggac ccaccttaga cttttaggat
7951 ccagaacaag gtgcagggtt tgccccaggc ctcaacatcc tgtcgcctgc
8001 cagctctcat atcctgctgg agaccaacaa gggccccagc ttcccaacag
8051 tcatggtaat cccagcagag atgctaaagg ggaaggaggc cccagggggc
8101 cgtgggctta ctggggctgg tgtctcccca cagGTGATAC AGATGTACAA
8151 GGAAAGCCCA GAGATCCTGA GTCCACCAGC TCCACAGGGA GCGGGGCTTA
8201 Ggtggggggc atcctcctcc ctgccccaga gctttgccag agtgactacc
8251 tccatgaccg tggccatcca cccaccccc ctgcctggtg cctacatcca
8301 tccagcccc ctgatgagccc cttggctccc tgctgtcact agctctggca
8351 acctcagttc caggggacca ggtccagcca ctgggtgaaa gagcagctga
8401 agcacagaga ccatgtgtgg ggcgtgtggg gtcactggga agcactgggt
8451 ctggtgttag acgcaggatg gacccctgga gggctctgct gctgctgcat
8501 cccctctccc gaccagctg tcatgggcct ccctgatatc catacagaac
8551 agccaccgat ttgcacatcc aggcctgtgt gagcctgtat ctgtgtcact
8601 tgagagtga aactggcact tggggctgca gtgcagccct gtcccccttc
8651 ccacccaca cactgcctg cccagctgac caagcctgag ggaccctcca
8701 gcacccttcc gtctgggtgac tcctgggcag gctctccata tccttgccca
8751 cctcctacca catccattat ttatatgaaa atgtctattt ttgtagtata
8801 catacatgtt agctatgatg aaagttttat tttttaaaga atgaaatata
8851 ttctatgtga agctatgatg aaagttttat tttttaaaga atgaaatata
8901 ttctatgtga actaatctcg aaagttttat tttttaaaga atgaaatata
8951 ttctatgtgt gcaagtgaac attagcttca gttgcttttt tttggacaga
9001 gtggggagtt tgcaagtga ccttagctat tggaaggagc ttctctgggtg
9051 ccaggacctg aggtattagc ttctctagtt ctgggtggaa aagacccag
9101 attctggatt tttgtcatat acttggtaac atcatctgga ttaagtgctt

```

Fig. 1 (cont.)

9151 actatacaaa acgataacaa attttgttgg tgtgaaatcc tactgggttc
 9201 aatctggaga ccgagagcag aaaaaaaga accccactgt gtggctttca
 9251 gagccaccat attccagcct gcccgctctt ccagactcac ctccacctac
 9301 ctgcttcacc cgcacgggaa acggcaaggc agagggggcaa agccatgcag
 9351 caggtggaag gcgaggtgga ggcagatcag gaaagcagcc agttgaagca
 9401 gagagaggtc aacaggggtct ggggagcttc tcaggaggtt tgtggaccca
 9451 gggaaaggag ccaggttcca gagcaacctc caaggcaaag gcctctgtaa
 9501 gttggttgtc ctgacagccg agaggtgtct ttggccagtc agccagtgga
 9551 tcagttgctg gaactgctca gaaactgagg tgctagcagt tagtgaggac
 9601 acagcgtaag ttgtttgttc tgtgaaagtt gaacagctcc actaagcaga
 9651 ggccttgaag agtggccaca gccctggaat agagcacaga gcctcaccta
 9701 gaggcgtggg gaggtttgca actgccccctt cccagccata gcttaggacc
 9751 catagtctag ttcacataga ccctgggctc caaccaccca ctcaccagga
 9801 atgatcccac cccaggaaca atgcgttctc acatcccacc ccacctggac
 9851 aaaggccagg aaatcatgtt ctgaccaaaa gatacaacaa caaaaacaac
 9901 aacaacaaaa aacgcctatt gcaattgaat ccacgctaaa atgcctaaaa
 9951 agctcaagag aagcgggtag ttggcagaga acctagagta gggggtgcaa
 10001 ccagcaggcc caaggagggg aggctgcatt tgggtccagc agtgtttggg
 10051 tcaccaagaa gggccttcta ggtggagcag agagagctca ccaggccaga
 10101 atagtgcaaa gggggtcagc cctcagtgcc acttaccagc ggagtaaccc
 10151 tgggeaagtt agccagcctc actaagcctc cccatcttca tctttccagG
 XON 22
 10201 **CCCGAGGAGA TC^ETAG^ECCCTCT GCCTCCCACC CCAGCACCCC CTCATCAGAC**
 10251 **ACAAGGAGCG CCACTGTCTG GACAGGCTGA ATTGGTCTTC GGGTCCCTAA**
 10301 **TTTCTCATAC GCCATTCCCT CTGCCTAGAA CACTTTCTCA CCTCCCCTTG**
 10351 **ATGTGACCCC ATATCACCCCT TCGAGGTGAA TTGGATCGGA TGCCATCTCC**
 10401 **TCCAGGAGGG GTGGGGTCGT GCCTCCTGTG AGGTCCCAGT GCCCCTGAGT**
 10451 **GTCTGTGCCC GTCTGTTTCC CCGTCCCTCT CTCTAAGCCC GGAGGCTTAC**

Fig. 1 (cont.)

```

10501  TCGGGTAAG GACGGCGGGA CAGGACCTTA ACCCCTGGGA CGAACACCAG
10551  CTCCGCAAAG GACTCCGCAC CCGGCGCCGC CCACGGGGTG CGGGTCCCAG
10601  GAGGACCAGC AGAGAGGAGC ATAGGAGAGC AAAGGAGATC AGTGACCCAT
10651  GGCTTCCCCG GTGGCGCGGA ACAGCCCGGA GCCGCCTGTG ATTTGCATAC
10701  CCATGGTGCA CCACGAAAAG ATACCCTCAA GATGCTTGCA CTCCCTCTGT
10751  GCGCGCATTT CTGCACTGTT TTAGAGCATG ATGCCTCTTA CACGCATCTG
10801  TGTGCATAAA CTACATATAG GGAGTGCGTA CCACGCAGGC ATCCAACAAC
10851  CATAAGTGTG TTAAGTGTTA GTTCTCCCTG CGAGGTTCGA AGCGGAAGTC
10901  ACGAATATAC TCGGGTTTCT CTTCAAAGCG CATAAATCTT TCGCCTTTTA
10951  CTAAAGATTT CCGTGGAGAG AAAGTTGTGA GTTTTTATTTC AATTTTTTGA
11001  GGCCTCTTAT TTCCTGAGGC TACATTTTTA AGTATTAAAA GTTAGGCAAC
11051  TACAAAAAAA AAAAAAAA

```

Fig. 1 (cont.)

```

1 .....MTRSPPLRELP..... 11
1. MASAGNAAEPQDRGGGSGCIGAPGRPAGGGRRRRTGGLRRAAAPDRDYL 50
12 ..PSYTPPARTAAPQI...LAGSLKAPLWLRAYFQGLLFSLGCGIQRHCG 56
51 HRPSYCDAA.FALEQISKGKATGRKAPLWLRKFQRLLPKLGCIYQKNCG 99
57 KVLFLGLLAFGALALGLRMAIIETNLEQLWVEVGSRVSQELHYTKEKLGE 106
100 KFLVVGLLIFGAFVGLKAANLETNVEELWVEVGGRVSRRELNYTRQKIGE 149
107 EAYTSQMLIQATARQEGENILTPEALGLHLQAALTASKVQVSLSYGKSWDL 156
150 EAMFNPQLMIQTPKEEGANVLTTEALLQHLDSALQASRVHVYMYNRQWKL 199
157 NKICYKSGVPLIENGMIERMIEKLFPCVILTPLDCEWEGAKLOGGSAYLP 206
200 EHLCKYSGELITETGYMDQIIIEYLYPCLIIITPLDCFWEGAKLQSGTAYLL 249
207 GRPDIQWTNLDPEQLLEELGPFA.SLEGFRELLDKAQVGQAYVGRPCLHP 255
250 GKPLRWTFNFDPLEFLEELKKINYQVDSWEEMLNKAEVGHGYMDRPLNLP 299
256 DDLHCPPSAPNHHSRQAPNVAHELSCGCHGFSHKFMHWQEELLLGGMARD 305
300 ADPDCPATAPNKNSTKPLDMALVLNGGCHGLSRKYMHWQEELIVGGTVKN 349
306 PQGELLRAEALQSTFLMSPROLYEHFRGDYQTHDIGWSEEQASTVLQAW 355
350 STGKLVSAAHALQTMFQLMTPKQMYEHFKGYEYVSHINWNEDKAAAILEAW 399
356 QRRFVQLAQEALPENASQOIHAFFSSTLDDILHAFSEVSAARVVGGYLLM 405
400 QRTYVEVVHQSV AQNSTQKVLSTTTTLDDILKSFSDVSVIRVASGYLLM 449
406 LAYACVTMLRWDCAQSGSVGLAGVLLVALAVASGLGLCALLGITFNAAT 455
450 LAYACLTMLRWDCSKSQGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAT 499
456 TQVLPFLALGIGVDDVFLLAHAFTEALPG..TPLQERMGECLQRTGTSVV 503
500 TQVLPFLALGVGVDDVFLLAHAFSETGQNKRIPFEDRTGECLKRTGASVA 549
504 LTSINNMAAFLMAALVPIPALRAFSLQAAIVVGCTFVAVMLVFPAILSLD 553
550 LTSISNVTAFFMAALIPIPALRAFSLQAAVVVFNFAMVLLIFPAILSMD 599
554 LRRRHCRQLDVLCCFSSPCSAQVIQILPQELGDGT.....VPVG 592
600 LYRREDRRLDIFCCFTSPCVSRVIQVEPQAYTDTHDNTRYSPPPPYSSHS 649
593 IAH.....LTATVQAFTHCEASSQHVVITILPPQAHL...VPPPSDPLGS 633
650 FAHETQITMQSTVQLRTEYDPHVVYTTAEPRSEISVQPVTVTQDTLSC 699

```

Fig. 2A

Fig. 2B Genómico

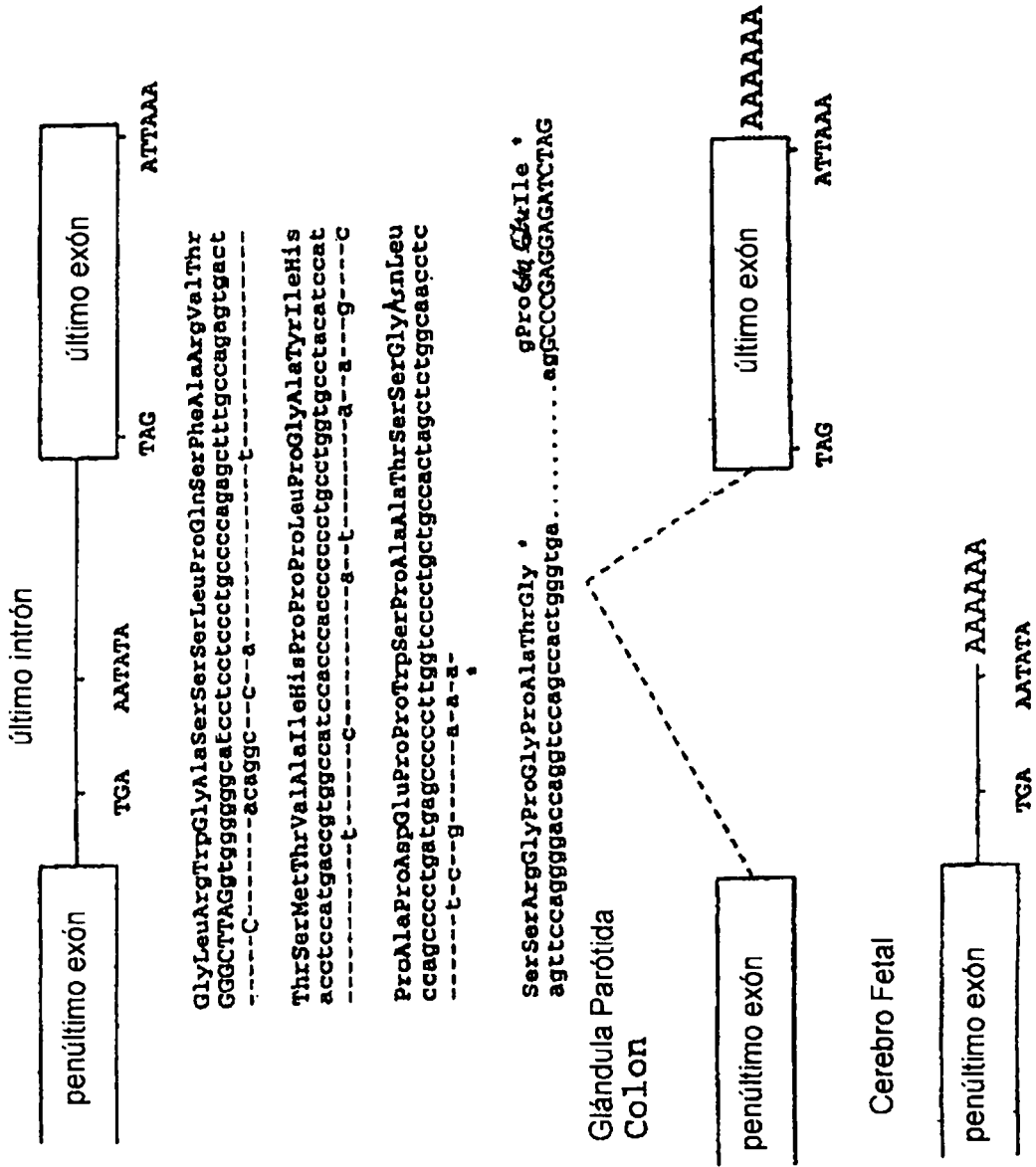
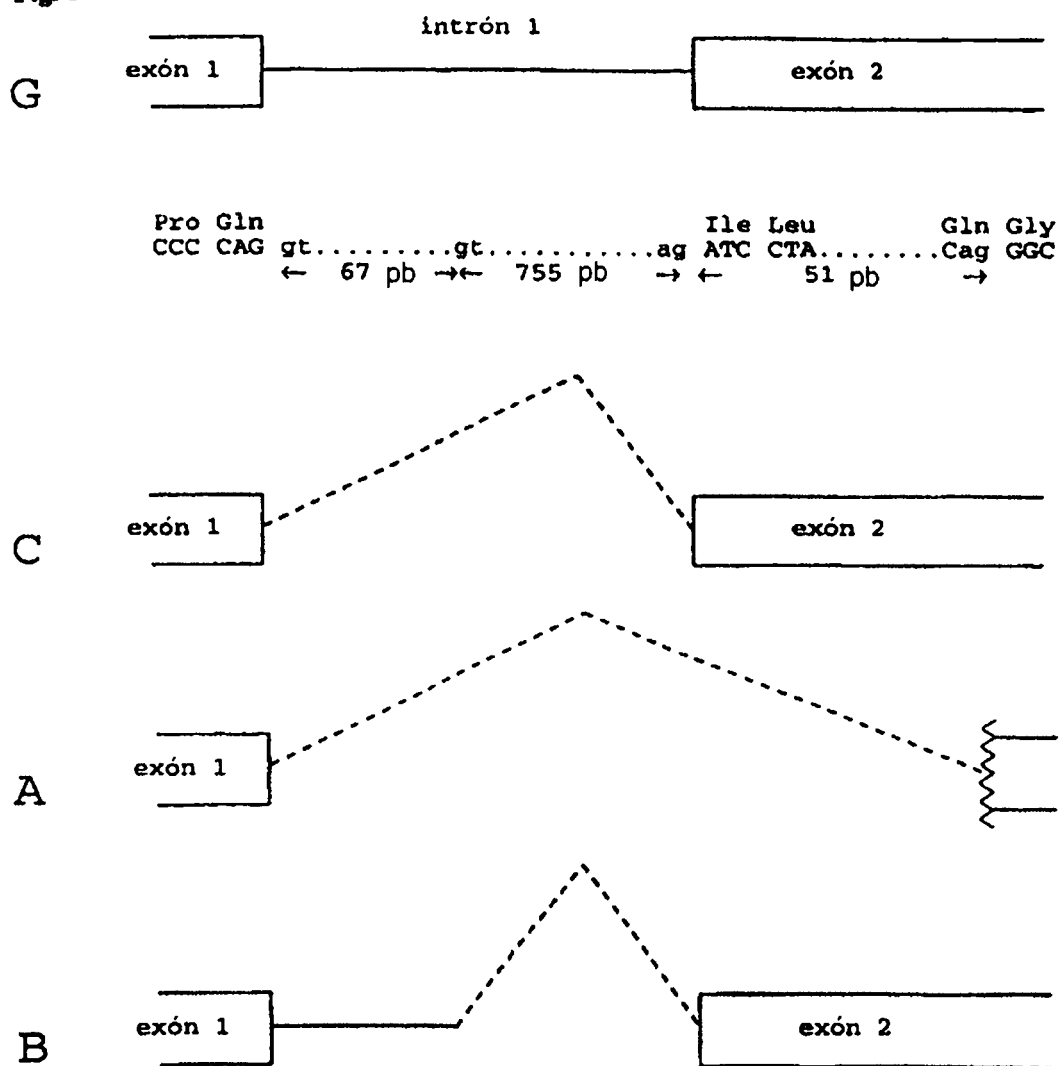


Fig. 2C



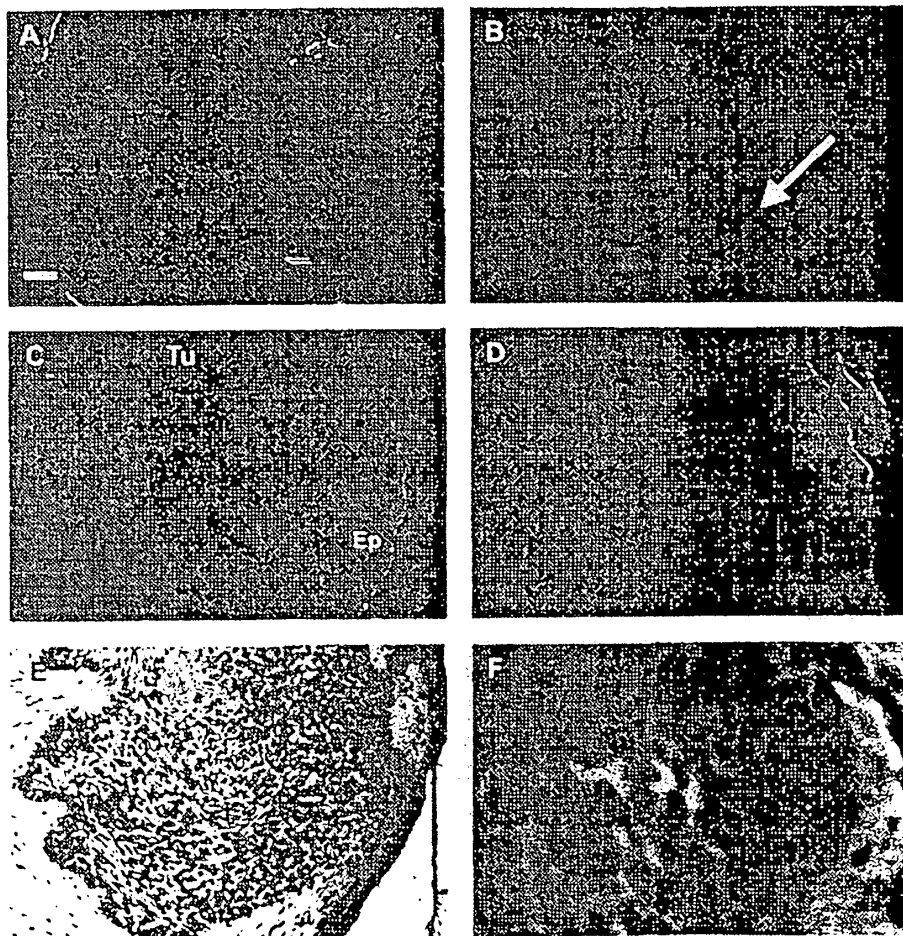


Fig. 3

ES 2 268 886 T3

LISTA DE SECUENCIAS

<110> PHARMACIA & UPJOHN S.P.A KAROLINSKA INNOVATIONS AB

5 <120> UN NUEVO COMPONENTE EN LA VÍA DE SEÑALIZACIÓN DE HEDGEHOG

<130> 56822

10 <140>

<141>

15 <160> 5

<170> PatentIn Ver. 2.1

<210> 1

20 <211> 1146

<212> PRT

<213> HUMANO

25 <400> 1

```

30      Met Thr Arg Ser Pro Pro Leu Arg Glu Leu Pro Pro Ser Tyr Thr Pro
          1          5          10          15

      Pro Ala Arg Thr Ala Ala Pro Gln Ile Leu Ala Gly Ser Leu Lys Ala
          20          25          30

35      Pro Leu Trp Leu Arg Ala Tyr Phe Gln Gly Leu Leu Phe Ser Leu Gly
          35          40          45

      Cys Gly Ile Gln Arg His Cys Gly Lys' Val Leu Phe Leu Gly Leu Leu
          50          55          60

40      Ala Phe Gly Ala Leu Ala Leu Gly Leu Arg Met Ala Ile Ile Glu Thr
          65          70          75          80

      Asn Leu Glu Gln Leu Trp Val Glu Val Gly Ser Arg Val Ser Gln Glu
          85          90          95

45      Leu His Tyr Thr Lys Glu Lys Leu Gly Glu Glu Ala Ala Tyr Thr Ser
          100          105          110

      Gln Met Leu Ile Gln Thr Ala Arg Gln Glu Gly Glu Asn Ile Leu Thr
          115          120          125

50      Pro Glu Ala Leu Gly Leu His Leu Gln Ala Ala Leu Thr Ala Ser Lys
          130          135          140

      Val Gln Val Ser Leu Tyr Gly Lys Ser Trp Asp Leu Asn Lys Ile Cys
          145          150          155          160

55      Tyr Lys Ser Gly Val Pro Leu Ile Glu Asn Gly Met Ile Glu Arg Met
          165          170          175

      Ile Glu Lys Leu Phe Pro Cys Val Ile Leu Thr Pro Leu Asp Cys Phe
          180          185          190

65      Trp Glu Gly Ala Lys Leu Gln Gly Gly Ser Ala Tyr Leu Pro Gly Arg
          195          200          205

```

ES 2 268 886 T3

Pro Asp Ile Gln Trp Thr Asn Leu Asp Pro Glu Gln Leu Leu Glu Glu
 210 215 220
 5 Leu Gly Pro Phe Ala Ser Leu Glu Gly Phe Arg Glu Leu Leu Asp Lys
 225 230 235 240
 Ala Gln Val Gly Gln Ala Tyr Val Gly Arg Pro Cys Leu His Pro Asp
 245 250 255
 10 Asp Leu His Cys Pro Pro Ser Ala Pro Asn His His Ser Arg Gln Ala
 260 265 270
 Pro Asn Val Ala His Glu Leu Ser Gly Gly Cys His Gly Phe Ser His
 275 280 285
 15 Lys Phe Met His Trp Gln Glu Glu Leu Leu Leu Gly Gly Met Ala Arg
 290 295 300
 20 Asp Pro Gln Gly Glu Leu Leu Arg Ala Glu Ala Leu Gln Ser Thr Phe
 305 310 315 320
 Leu Leu Met Ser Pro Arg Gln Leu Tyr Glu His Phe Arg Gly Asp Tyr
 325 330 335
 25 Gln Thr His Asp Ile Gly Trp Ser Glu Glu Gln Ala Ser Thr Val Leu
 340 345 350
 Gln Ala Trp Gln Arg Arg Phe Val Gln Leu Ala Gln Glu Ala Leu Pro
 355 360 365
 30 Glu Asn Ala Ser Gln Gln Ile His Ala Phe Ser Ser Thr Thr Leu Asp
 370 375 380
 35 Asp Ile Leu His Ala Phe Ser Glu Val Ser Ala Ala Arg Val Val Gly
 385 390 395 400
 Gly Tyr Leu Leu Met Leu Ala Tyr Ala Cys Val Thr Met Leu Arg Trp
 405 410 415
 40 Asp Cys Ala Gln Ser Gln Gly Ser Val Gly Leu Ala Gly Val Leu Leu
 420 425 430
 Val Ala Leu Ala Val Ala Ser Gly Leu Gly Leu Cys Ala-Leu-Leu Gly-
 435 440 445
 45 Ile Thr Phe Asn Ala Ala Thr Thr Gln Val Leu Pro Phe Leu Ala Leu
 450 455 460
 Gly Ile Gly Val Asp Asp Val Phe Leu Leu Ala His Ala Phe Thr Glu
 465 470 475 480
 50 Ala Leu Pro Gly Thr Pro Leu Gln Glu Arg Met Gly Glu Cys Leu Gln
 485 490 495
 55 Arg Thr Gly Thr Ser Val Val Leu Thr Ser Ile Asn Asn Met Ala Ala
 500 505 510
 Phe Leu Met Ala Ala Leu Val Pro Ile Pro Ala Leu Arg Ala Phe Ser
 515 520 525
 60 Leu Gln Ala Ala Ile Val Val Gly Cys Thr Phe Val Ala Val Met Leu
 65

ES 2 268 886 T3

	530	535	540
5	Val Phe Pro Ala Ile Leu Ser Leu Asp Leu Arg Arg Arg His Cys Gln 545 550 555 560		
	Arg Leu Asp Val Leu Cys Cys Phe Ser Ser Pro Cys Ser Ala Gln Val 565 570 575		
10	Ile Gln Ile Leu Pro Gln Glu Leu Gly Asp Gly Thr Val Pro Val Gly 580 585 590		
	Ile Ala His Leu Thr Ala Thr Val Gln Ala Phe Thr His Cys Glu Ala 595 600 605		
15	Ser Ser Gln His Val Val Thr Ile Leu Pro Pro Gln Ala His Leu Val 610 615 620		
	Pro Pro Pro Ser Asp Pro Leu Gly Ser Glu Leu Phe Ser Pro Gly Gly 625 630 635 640		
20	Ser Thr Arg Asp Leu Leu Gly Gln Glu Glu Glu Thr Arg Gln Lys Ala 645 650 655		
	Ala Cys Lys Ser Leu Pro Cys Ala Arg Trp Asn Leu Ala His Phe Ala 660 665 670		
25	Arg Tyr Gln Phe Ala Pro Leu Leu Leu Gln Ser His Ala Lys Ala Ile 675 680 685		
30	Val Leu Val Leu Phe Gly Ala Leu Leu Gly Leu Ser Leu Tyr Gly Ala 690 695 700		
	Thr Leu Val Gln Asp Gly Leu Ala Leu Thr Asp Val Val Pro Arg Gly 705 710 715 720		
35	Thr Lys Glu His Ala Phe Leu Ser Ala Gln Leu Arg Tyr Phe Ser Leu 725 730 735		
	Tyr Glu Val Ala Leu Val Thr Gln Gly Gly Phe Asp Tyr Ala His Ser 740 745 750		
40	Gln Arg Ala Leu Phe Asp Leu His Gln Arg Phe Ser Ser Leu Lys Ala 755 760 765		
	Val Leu Pro Pro Pro Ala Thr Gln Ala Pro Arg Thr Trp Leu His Tyr 770 775 780		
45	Tyr Arg Asn Trp Leu Gln Gly Ile Gln Ala Ala Phe Asp Gln Asp Trp 785 790 795 800		
50	Ala Ser Gly Arg Ile Thr Arg His Ser Tyr Arg Asn Gly Ser Glu Asp 805 810 815		
	Gly Ala Leu Ala Tyr Lys Leu Leu Ile Gln Thr Gly Asp Ala Gln Glu 820 825 830		
55	Leu Leu Asp Phe Ser Gln Leu Thr Thr Arg Lys Leu Val Asp Arg Glu 835 840 845		
60	Gly Leu Ile Pro Pro Glu Leu Phe Tyr Met Gly Leu Thr Val Trp Val 850 855 860		

65

ES 2 268 886 T3

Ser Ser Asp Pro Leu Gly Leu Ala Ala Ser Gln Ala Asn Phe Tyr Pro
 865 870 875 880
 Pro Pro Pro Glu Trp Leu His Asp Lys Tyr Asp Thr Thr Gly Glu Asn
 885 890 895
 Phe Arg Ile Pro Pro Ala Gln Pro Leu Glu Phe Ala Gln Phe Pro Phe
 900 905 910
 Leu Leu Arg Gly Leu Gln Lys Thr Ala Asp Phe Val Glu Ala Ile Glu
 915 920 925
 Gly Ala Arg Ala Ala Cys Ala Glu Ala Gly Gln Ala Gly Val His Ala
 930 935 940
 Tyr Pro Ser Gly Ser Pro Phe Leu Phe Trp Glu Gln Tyr Leu Gly Leu
 945 950 955 960
 Arg Arg Cys Phe Leu Leu Ala Val Cys Ile Leu Leu Val Cys Thr Phe
 965 970 975
 Leu Val Cys Ala Leu Leu Leu Leu Asn Pro Trp Thr Ala Gly Leu Ile
 980 985 990
 Val Leu Val Leu Ala Met Met Thr Val Glu Leu Phe Gly Ile Met Gly
 995 1000 1005
 Phe Leu Gly Ile Lys Leu Ser Ala Ile Pro Val Val Ile Leu Val Ala
 1010 1015 1020
 Ser Val Gly Ile Gly Val Glu Phe Thr Val His Val Ala Leu Gly Phe
 1025 1030 1035 1040
 Leu Thr Thr Gln Gly Ser Arg Asn Leu Arg Ala Ala His Ala Leu Glu
 1045 1050 1055
 His Thr Phe Ala Pro Val Thr Asp Gly Ala Ile Ser Thr Leu Leu Gly
 1060 1065 1070
 Leu Leu Met Leu Ala Gly Ser His Phe Asp Phe Ile Val Arg Tyr Phe
 1075 1080 1085
 Phe Ala Ala Leu Thr Val Leu Thr Leu Leu Gly Leu Leu His Gly Leu
 1090 1095 1100
 Val Leu Leu Pro Val Leu Leu Ser Ile Leu Gly Pro Pro Pro Glu Val
 1105 1110 1115 1120
 Ile Gln Met Tyr Lys Glu Ser Pro Glu Ile Leu Ser Pro Pro Ala Pro
 1125 1130 1135
 Gln Gly Gly Gly Leu Arg Pro Glu Glu Ile
 1140 1145

60 <210> 2
 <211> 3453
 <212> ADN
 65 <213> HUMANO

ES 2 268 886 T3

<400> 2

	atccccgcca	gcatgactcg	atcgccgccc	ctcagagagc	tgcccccgag	ttacacacccc	60
	ccagctcgaa	ccgcagcacc	ccagatccta	gctgggagcc	tgaaggctcc	actctggctt	120
5	cgtgcttact	tccagggcct	gctcttctct	ctgggatgcg	ggatccagag	acattgtggc	180
	aaagtgtctt	ttctgggact	gttggccttt	ggggccctgg	cattaggtct	ccgcatggcc	240
	attattgaga	caaacttggg	acagctctgg	gtagaagtgg	gcagccgggt	gagccaggag	300
	ctgcattaca	ccaaggagaa	gctgggggag	gaggctgcat	acacctctca	gatgctgata	360
10	cagaccgcac	gccaggaggg	agagaacatc	ctcacacccg	aagcaacttg	cctccacctc	420
	caggcagccc	tcactgccag	taaagtccaa	gtatcactct	atgggaagtc	ctgggatttg	480
	aacaaaatct	gctacaagtc	aggagtcccc	cttattgaaa	atggaaatgat	tgagcggatg	540
	attgagaagc	tgtttccgtg	cgatgacctc	acccccctcg	actgcttctg	ggaggggagcc	600
	aaactccaag	ggggctccgc	ctacctgccc	ggccgcccgg	atatccagtg	gaccaacctg	660
15	gatccagagc	agctgctgga	ggagctgggt	ccctttgcct	cccttgaggg	cttccgggag	720
	ctgctagaca	aggcacaggt	gggccaggcc	tacgtggggc	ggccctgtct	gcacctgat	780
	gacctccact	gcccacctag	tgcccccaac	catcacagca	ggcaggctcc	caatgtggct	840
	cacgagctga	gtgggggctg	ccatggcttc	tcccacaaat	tcatgcaactg	gcaggaggaa	900
	ttgctgctgg	gaggcatggc	cagagacccc	caaggagagc	tgctgagggc	agaggccctg	960
20	cagagcacct	tcttgctgat	gagtcctcgc	cagctgtacg	agcatttccg	gggtgactat	1020
	cagacacatg	acattggctg	gagtggaggag	caggccagca	cagtgtctaca	agcctggcag	1080
	cggcgctttg	tgcagctggc	ccaggaggcc	ctgcctgaga	acgcttccca	gcagatccat	1140
	gccttctctc	ccaccaccct	ggatgacatc	ctgcatgcgt	tctctgaagt	cagtgtctgc	1200
	cgtgtggtgg	gaggctatct	gctcatgctg	gcctatgcct	gtgtgacctat	gctgcgggtg	1260
25	gactgcgccc	agtcccaggg	ttccgtgggc	cttgccgggg	tactgtctgt	ggccctggcg	1320
	gtggcctcag	gccttgggct	ctgtgccctg	ctcggcatca	ccttcaatgc	tgccactacc	1380
	caggtgtctg	ccttcttggc	tctgggaatc	ggcgtggatg	acgtattcct	gctggcgcat	1440
	gccttccacg	aggctctgcc	tggcacccct	ctccaggagc	gcattgggca	gtgtctgcag	1500
	cgcacgggca	ccagtgtcgt	actcacatcc	atcaacaaca	tggccgcctt	ctcatgtgct	1560
30	gcctctgttc	ccatccctgc	gctgcgagcc	ttctccctac	aggcgcccat	agtgggttgg	1620
	tgcacctttg	tagccgtgat	gcttgtcttc	ccagccatcc	tcagcctgga	cctacggcgg	1680
	cgcactgpc	agcgcttga	tgtgctctgc	tgcttctcca	gtccctgtct	tgctcaggtg	1740
	attcagatcc	tgccccaggga	gctgggggac	gggacagtac	cagtgggcat	tgccacctc	1800
35	actgccacag	ttcaagcctt	taccactgt	gaagccagca	gccagcatgt	ggtcaccatc	1860
	ctgcctcccc	aagcccaacct	ggtgccccca	ccttctgacc	cactgggctc	tgagctcttc	1920
	agccctggag	ggtccacacg	ggaccttcta	ggcaggagg	aggagacaag	gcagaaggca	1980
	gcctgtcaagt	ccctgcccctg	tgcccgtg	aatcttgccc	atttcgccc	ctatcagttt	2040
	gccccgttgc	tgctccagtc	acatgctaag	gccatcgtgc	tggtgtctct	tggtgtctct	2100
40	ctgggcctga	gcctctacgg	agccaccttg	gtgcaagacg	gcctggccct	gacggatgtg	2160
	gtgcctcggg	gcaccaaggga	gcatgccttc	ctgagcgccc	agctcaggta	cttctccctg	2220
	tacgaggtgg	ccctgggtgac	ccagggtggc	tttgactacg	cccaactcca	acgcgcctc	2280
	tttgatctgc	accagcgctt	cagttccctc	aaggcggtgc	tgccccccac	ggccacccag	2340
	gcaccccgca	cctggctgca	ctattaccgc	aactggctac	agggaatcca	ggctgccttt	2400
45	gaccaggact	gggcttctgg	gcgcataccc	cgccactcgt	accgcaatgg	ctctgaggat	2460
	ggggccctgg	cctacaagct	gctcatccag	actggagagc	cccaggagct	tctggatttc	2520
	agccagctga	ccacaaggaa	gctgggtggc	agagagggac	tgattccacc	cgagctcttc	2580
	tacatggggc	tgaccgtgtg	ggtgagcagt	gacccctggg	gtctggcagc	ctcacaggcc	2640
	aacttctacc	ccccacctcc	tgaatggctg	cacgacaaat	acgacaccac	gggggagaa	2700
50	tttcgcatcc	cgcagctca	gcccttggag	tttgcccagt	tcccttctct	gctgcgtggc	2760
	ctccagaaga	ctgcagactt	tgtggaggcc	atcgaggggg	cccgggcagc	atgcgcagag	2820
	gccggccagg	ctgggggtgca	cgcctacccc	agcggtctcc	ccttctctct	ctgggaacag	2880
	tatctggggc	tgccggcgtg	cttctgtctg	gccgtctgca	tctgtctggt	gtgcacttct	2940
	ctcgtctgtg	ctctgtctgt	cctcaacccc	tggacggctg	gcctcatagt	gctggctcct	3000
55	gcgatgatga	cagtggaaact	ctttggtatc	atgggtttcc	tgggcatcaa	gctgagtgcc	3060
	atccccgtgg	tgatccttgt	ggcctctgta	ggcatgtggc	ttgagttcac	agtccactgt	3120
	gctctgggct	tcctgaccac	ccagggcagc	cggaaacctgc	gggcccgcga	tgcccttgag	3180
	cacacatttg	cccccgtagc	cgatggggcc	atctccacat	tgctgggtct	gctcatgctt	3240
60	gctgggtccc	actttgactt	cattgtaagg	tacttctttg	cggcgctgac	agtgtctcac	3300
	ctcctggggc	tcctccatgg	actcgtgctg	ctgcctgtgc	tgctgtccat	cctgggccc	3360
	ccgccagagg	tgatacagat	gtacaaggaa	agcccagaga	tcctgagtcc	accagctcca	3420
	cagggaggcg	ggcttagggc	caggagatc	tag			3453

<210> 3

<211> 126

ES 2 268 886 T3

<212> ADN

<213> HUMANO

5 <400> 3

```
gagcgcatgg gcgagtgtct gcagcgcacg ggcaccagtg tcgtactcac atccatcaac 60
aacatggccg ccttcctcat ggctgccctc gttcccatcc ctgcgctgcg agccttctcc 120
ctacag 126
```

10

<210> 4

<211> 119

<212> ADN

15

<213> HUMANO

<400> 4

```
gcggcccatag tgggttgctg cacctttgta gccgtgatgc ttgtcttccc agccatcctc 60
agcctggacc tacggcggcg ccactgccag cgccttgatg tgctctgctg cttctccag 119
```

<210> 5

25

<211> 12886

<212> ADN

<213> HUMANO

30

<400> 5

```
cgggtgaatc ccggcgccgc gccccggacc cgcagctccc tgcactcctc cctcccagcc 60
gctttaacac ccacacccca cagtctctcc cacgscgcg ccttggcgcg cccactgaat 120
ccctacgcgg ggcccagcgg tacccggaga ccgggctagc ctatgggagc gcccagataa 180
cgcgggttg gggcgcccgc gccccatcc ccgccagcat gactcgatcg ccgcccctca 240
gagagctgcc cccgagttac acacccccag ctcgaaaccg agcaccacag gtgagtagag 300
ggggagctgg aagaaggaa agagcggagc caggtctgtc actcgggcct ctgcaagggt 360
tgtgatgtct tgaagtgcg agtgtcatta gatgtctgaa ggcaagttag agccagcacc 420
gcaagcaagt tgtgcgtgtg tgtcgggtgt tctgtgccgg tgtctcctca tcgtctggcc 480
agtgagaatg aatgtctgtg gggtcacctc tgtgtccacc cgacgacagg tgtgtgtaca 540
tatgtatcct gctctcagaa aatgggccta tgccgcggg cgcggtgact cagcctgta 600
atcccaacac tgggaggctg aggcaggcag attacctgag gtcaggagtt cgagaccagc 660
caggccaaca tggggaaact ctgtctctac taaaaataaa aattagcagg gcgtgggtgg 720
gggcgcctgt agtcccaact actcgggagg ctgaggcagg agaattctct gaacctggga 780
ggcggaggtt gcagtcaagc cgagatcaca ccactgcact ccagccaggg caacagagcg 840
agatgcgtct caaaaaaaaa aaaaaaaaaa aaaaggagag aaacaaaaaa gaaaagaaag 900
gaaaataggc ctatgccttc ctgaggtgtg tgctggggat ggtgggtgtt acatcttcca 960
agtctgggcc tgtgtctgtg ttggtgctcc ctgtcccaca tccagaaatc aagaagcgag 1020
ggctgggcag cagatatata gggtgagaag ggaaggattt catgcattgt tacagtgatg 1080
cctggctgac ccttctcttt ccatcccaga tcctagctgg gagcctgaag gctccactct 1140
ggcttcgtgc ttacttccag ggctgtctt tctctctggg atgcgggatc cagagacatt 1200
gtggcaaatg gctctttctg ggactgttgg cctttggggc cctggcatta ggtctccgca 1260
tggccattat tgagacaaac ttggaacagc tctgggtaga aggtaagttg tggacactgg 1320
ccatagctgc tcagggtatg tgagcccaag acaagaacgg ggtgaggagc tggctactga 1380
gctctagcag gcctggccct ggggtgag ggtcacctgc tggccagagt cctacacctg 1440
ggcatctagg aacaccgttc cctcccacaa ccataaatgg acatctgcag atgttagacc 1500
ttgtgaattg ggtgtataga atatttgtct gtgattttgc ttaagcccta gaccacacaa 1560
gaaatccgca gagccacaca accgaggctt ctccacttcc ctggcagagc tccaggaccc 1620
agaagcctgg ggagtcggcc ttggccatga ccaggccctc agggataagg cagggactcc 1680
tctctgccgc acctctccca cttgtctggg gaggcctcag agttagaagc ccccttcgca 1740
agttctgacc cgcctaggcc ctttcccccc cagctttctc tgtttgctct ggcagtgggc 1800
agccgggtga gccaggagct gcattacacc aaggagaagc tgggggagga ggctgcatac 1860
acctctcaga tgctgatata gaccgcacgc caggaggagg agaacatcct cacacccgaa 1920
gcacttggcc tccacctcca ggcagccctc actgccagta aagtccaagt atcactctat 1980
```

65

ES 2 268 886 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

ggggaaggtga gtctggctga gccctgagc agctgggggc gaggcgtgct gtggggggttc 2040
 tggagtggga atcccccttct tctgctgac tcctatgccc ctggctattg cagtcctggg 2100
 atttgaacaa aatctgctac aagtcaggag tcccccttat tgaaaatgga atgattgagc 2160
 gggtaagtgt cctgagaggg agtagaggca gaactttttc tgtagcgtgg gaggactcag 2220
 agaccgagca agccccacag cctgcaatct gcccccttaa aactaaggag ggggattgca 2280
 gagggcatcc taaaaaggtt gtggggcagg actgacgtgg cccgggggat ccctggcaga 2340
 tgattgagaa gctgtttccg tgcgtgatcc tcacccccct cgactgcttc tgggaggag 2400
 ccaaactcca agggggctcc gccctacctg cgtgagtgcc actcctgggg ccctgcttca 2460
 tctcccgctg gggactctcc cagcagaaag gaggggtctg gggaaatgagg atgatcaaaa 2520
 ccttaccaag gtcctaatta cctcccaggc caggaacaga gagcatgggc tcccccaagg 2580
 ctctctccac atcctccttc tctttccctc tcaagggaagg aagacctgac ttattttacac 2640
 aaaactaaac acaaagatct gtaagatctg agcaaaggag aaaaagatcc ccacaaagag 2700
 gctttgctgg gggaaattac ctagggtgtt gctaagccat tgcccaggcc agaaagaaaa 2760
 cctgtacag gcatgtgcct gctggttgta tattagaacc aagcacacag atgttaagg 2820
 aactcagtg ggcctttctg ggccctttct atgtattagg taaccctgcc ctgatattcg 2880
 tctcagcccc ttgtactctt ctacagctca ctgtagcacc ctggtgggcc catgcagcct 2940
 ggcagttctg agaagctgag gcttgacac cctccatatg gaaggacaaa tggcgagata 3000
 agaggagggt ggggtacagc atggcgcccc agcagcagtt tggagcctgg gttttcgtcc 3060
 ctgacctca ccaactatag gcttttccct cagcgccgc ccggatatcc agtggacca 3120
 cctggatcca gagcagctgc tggaggagct gggteccctt gcctcccttg agggcttccg 3180
 ggagctgcta gacaaggcac aggtgggcca ggcctacgtg gggcgccct gtctgcacc 3240
 tgatgacctc cactgccac ctagtcccc caaccatcac agcaggcagg tgggttccaa 3300
 ccaggtctgc cagggaaagg ctgttttctt tccctttccc ttcctcatat tcctgtgttc 3360
 tgggggagct gactgctctg tgccctgacc ccccaacttc tggccattat taccctgtc 3420
 ccacagtgcc agggcccca tggtccattc ccattcagtt atcctacgga gccctcaagt 3480
 ggtatatatg aatccctttt tccctttcta agcctagata aggctggact tctttttttt 3540
 tttttttttg agtctcactc tgtcacccag gctggagtgc agtagttcga tcttggctca 3600
 ctgcaacctc ggctcaagca attctcctgc cttagcctcc tgagtactg ggattacag 3660
 tgccaccac cctgcccggc taatttttat aagtgctggg attacaggcg 3720
 tgagccactg cgectggcca aggctggact ttttatcaaa atagactaat acagggaac 3780
 taagaacaca gcaggtaac atgaatatca taccctggtt cccaggttcc tttgtggccc 3840
 tgcaaatgtg gtactttttt cagaatccgc cagttacacc agctcctccc agaagcctac 3900
 ttccaggcct ctgcttcccc ttggggcttc ctgtctgcgg gatactagct gttcactcct 3960
 gcagagcagt caagaggctc agaatagtta cctacactcc agccctactg agcttcatgg 4020
 cagcgtggtt cctggaggtg gaagcccagg gacactcagt tatccacggc cagggccttg 4080
 agcattaacc cctcctgttc cctccagggt ctcccaatgt ggctcacgag ctgagtgggg 4140
 gctgccatgg ctctccccc aaattcatgc actggcagga ggaattgctg ctgggaggca 4200
 tggccagaga cccccaagga gagctgctga ggtagggtct cctctgggag ttggtgagg 4260
 gactctgttc atgagaacct atactgtaat gccaggcagc tctggcaaaa ggccctcac 4320
 atccctcacc aggtgttttg gccagctctg acccctggtt ctccacacc cccaccaggg 4380
 cagaggccct gcagagcacc ttcttgctga tgagtcccc ccagctgtac gacatttcc 4440
 ggggtgacta tcagacacat gacattggct ggagtggga gcaggccagc acagtgtac 4500
 aagcctggca gcggcgcttt gtgcagctcg gtatggacaa ggacaagggg ggtgccctga 4560
 ggccattccc tccctctgcc cctcctatc caccctgttt ctccagctgg cccaggaggc 4620
 cctgctgag aacgcttccc agcagatcca tgccttctcc tccaccacc tggatgacat 4680
 cctgcatgct ttctctgaag tcagtgtgc ccgtgtggtg ggaggctatc tgctcatggt 4740
 gggctcttgca cctggcacct tgeccccacc ccacctcaa ccagtgccca ccctgggagc 4800
 ccctgagact gccctttccc cccacagctg gcttatgcct gtgtgacct gctgcggtg 4860
 gactgcgccc agtcccaggg ttccgtgggc cctgcccggg tactgctggt ggccctggcg 4920
 gtggcctcag gccttgggct ctgtgccctg ctggcatca ccttcaatgc tgccactacc 4980
 caggtacgcc aggactgcag ggcagactca gtgccagtca ccaggcttca cgggtcctca 5040
 gctgcccgt cctctgcccc tccagtgct gcccttcttg gctctgggaa tggcggtgga 5100
 tgacgtattc ctgctggcgc atgecttcc agaggctctg cctggcacc ctctccagg 5160
 ggggccttgt cccccagggc tcatctgagg cagctcagct tactggttaa gagcctcttg 5220
 gttcaagtga cccttgggct gctaataaac ctgggtgcct cttgtcccca tctgtaaaaa 5280
 ggggaaataa taagtctgtg tcctaagggt tattgtttgg atcagtgagg taactcaagt 5340
 tgaatgctta gaacagccca tcatacgtac atggtaccca ataaatgcta gccactgtgt 5400
 tatgactgcc ccacctctgc accccaagtt cctgagcctc cccttcaact cactttgaca 5460
 cggccctccc cttgtgacct gagggcagg cccactctg tccctggcagg agcgcaggg 5520
 cgagtgtctg cagcgcacgg gcaccaagtgt cgtactcaca tccatcaaca acatggccgc 5580
 cttcctcatg gctgccctcg ttcccatccc tgcgtgctga gccttctccc tacaggcggc 5640

ES 2 268 886 T3

5 catagtgggt ggctgcacct ttgtagccgt gatgcttgte tccccagcca tcctcagcct 5700
 ggacctacgg cggcgccact gccagcgcc tgatgtgctc tgctgcttct ccaggtactg 5760
 cgtgcgcccc agccccctcc tcccgtagcc cagccagcc tgctccctca ccagcatttc 5820
 aaggcacaga cctgtcatcc actctctacc tcttccagtc cctgctctgc tcaggtgatt 5880
 cagatcctgc ccaggagct gggggacggg acagtaccag tgggcattgc ccacctcact 5940
 gccacagttc aagcctttac ccactgtgaa gccagcagcc agcatgtggt caccatcctg 6000
 cctccccaag cccacctggt gccccacct tctgaccac tgggctctga gctcttcagc 6060
 cctggagggt ccacacggga ccttctaggg caggaggagg agacaaggca gaaggcagcc 6120
 10 tgcaagtccc tgccctgtgc ccgctggaat cttgccatt tcgcccgtca tcagtttgcc 6180
 ccgttgctgc tccagtcaca tgctaaggta agactgggca gagcagggca gagacttagc 6240
 atctctgggc ccagaagggc agagagggt tagtccactg cctgaggggc tgggggcagc 6300
 cctgggggtc ccagcttagt tgctacatcc cgcaggccat cgtgctggtg ctctttggtg 6360
 ctctctggg cctgagcctc tacggagcca ccttgggtgca agacggcctg gccctgacgg 6420
 15 atgtggtgcc tcggggcacc aaggagcatg ccttctctgag cgcccagctc aggtacttct 6480
 ccctgtacga ggtggccctg gtgaccagg gtggctttga ctacggccac tcccaacggc 6540
 cectctttga tctgcaccag cgcttcagtt ccctcaaggc ggtgctgccc ccaccggcca 6600
 cccaggcacc ccgcacctgg ctgcaactat accgcaactg gctacagggt gagaggcgag 6660
 gagacgggca gggagggggt ctgcaaggag aaacgccctg gggccaccag ctaatagaac 6720
 20 cctatcctgg tctccccag gaatccaggc tgcccttgac caggactggg ctcttggcg 6780
 catcacccgc cactcgaccg caatggctct gaggatggg ccctggccta caagctgctc 6840
 atccagactg gagacgcca ggagcttctg gatttcagcc aggttgggag agggctggag 6900
 gggctccacta gtacaggggc tgacggcctc ctggggccag gccttcagcc ctctctgct 6960
 ctgcagctga ccacaaggaa gctggtggac agagagggac tgattccacc cgagctcttc 7020
 25 tacatggggc tgaccgtgtg ggtgagcagt gacccctgg gtctggcagc ctcacaggcc 7080
 aacttctacc cccacacctc tgaatggctg cagacaaat acgacaccac gggggagaac 7140
 tttcgagtg agtctgggg ggaactcgcc aagagcctca gcctcgcca cacaagccct 7200
 gagcctgagg ccctgcccac tctgcccctg gctcacccgc ctgtccctct cctcttctc 7260
 ccttccccct cctccacag tcccggcagc tcagcccttg gaggttgccc agttcccttt 7320
 30 cctgctgctg ggccctccaga agactgcaga ctttgtggag gccatcgagg gggccgggc 7380
 agcatgcgca gaggcgggac aggtgggggt gcacgcctac ccagcggt ccccttccct 7440
 ctctctggga cagtatctgg gcctgcggcg ctgcttctct ctggccgtct gcactctgct 7500
 ggtgtgcact tctctgctc gtgctctgct gctcctcaac ccctggacgg ctggcctcat 7560
 35 agtgagtgtc tgaggagtg gggacagaga gtctccctca ggtgctggtc ctggcgatga 7620
 ccctcctgcc agggagccctc tgtgagccct gtctccctca ggtgctggtc ctggcgatga 7680
 tgacagtggg actctttggt atcatgggtt tccctgggcat caagctgagt gccatccccg 7740
 tgggtatcct tgtggcctct gtaggcattg gcgttgagtt cacagtccac gtggctctgg 7800
 tgagcacggg caccctgggg agggaccaat cagctgattc agtattcaac acatattggt 7860
 caagccccta ctatgtgcta ggtactattt aagaatttgg gctgggtgga cgtggtagct 7920
 40 catctctgta atccagcac ttggggaggc caggagcagg ggtacacctg aggtcaggag 7980
 ttcgaaacca gcctggccaa catggtgaaa ccctgtcttt actaaaaata caaaaaatta 8040
 gccaggcggt gtggcacatg ccagtaatcc cagctacttt ggaggctgag gcagaattgc 8100
 ttgaacctgg gaggcgaagg ttgcagttag ctgagatcgt gccattgcac tccagcctgg 8160
 gcaacaagag tgaactctc cgtctcaaaa aaaaaaaaaa aaagaattt gggctgggca 8220
 45 cagtggctca tgccgtgta tgggatgat ctggggcatt ttgggaggcc aaggcaggcg 8280
 gatccccgta agtcaggagt tcaagaccag cctggccaac attgcaaac cccgtctcta 8340
 ctgaaaatac aaaaattagc tgggcgtggt ggctcatgcc tgtaatccca gctactcagg 8400
 aggctgaggc aggagaatta cttgaacca ggaggcgagg gttacagtga gctgagatca 8460
 catcactgta ctccagcctg ggcaaaagag caagattcaa tcttcaaaa aagaatttgg 8520
 50 aaaaataaaa taanaagaat acgggatata atagcaatac agttttttac ctccaaggaa 8580
 cttatattct agggatagag atagacaata agggctgggt gaggtggctt acgcctgtaa 8640
 tcccagcact ttgggaagcc gaggtgggca catcacttga ggtcaggagt tcaagaccag 8700
 cctggccgac atggtgaaac cccatctcta ctaaaaatat gaaaattagc tgggtgtggg 8760
 ggtgcatgcc tgtaatccca gctacttggg aggtcagggc aggagaatca cttgaaccgg 8820
 55 ggagggtgga ggttgccatg agccgagacc atgccactgc actccagcct gggagaccga 8880
 gcaaaactcc atctcaaaa aagaaaaaaa aacctcagcc tcccaaagtg ctgggattac 8940
 aggcattgag caccgtgccc gggcttgttt gttttaagag acagggtctc actctgtcac 9000
 ccaggctgta ccctattttc tacgtgtctc tgtgtctaag ctcaaaaac ccatccatac 9060
 agtcttacct aaaaaggact acgcatgct cctgtacagt tcccatgaac ttagttcacc 9120
 60 cagccttttg tgggcaagg tatcgttag ctctgaagac agatacagct cgggacatgg 9180
 ctgagctggc catgactggc agaggagcag ctccaggacc actctgtttt cctagggctt 9240
 cctgaccacc cagggcagcc ggaacctgcg ggccgcccac gcccttgagc acacatttgc 9300

ES 2 268 886 T3

ccccggtgacc gatgggggcca tctccacatt gctgggtctg ctcatgcttg ctggttccca 9360
 ctttgacttc attgtaaggt agggagggct cggggcaggg aggcagggct caggacaggc 9420
 ctgggctgac tccccccaca ccctaccctt aggtacttct ttgcgcgct gacagtgtct 9480
 5 acgctcctgg gcctcctcca tggactcgtg ctgctgctgt tgctgctgtc catcctgggc 9540
 ccgcccgcag aggtgaccac accctcggca ccatccctct actcccagcc caagggaagg 9600
 ggtagggaga ggcaaggga tggcagcctc cctgtggccc acagacaggt acctcccca 9660
 caggtgccac cagctgaagg tggcagcctc cctcttccc cagacacat gttcctgccc 9720
 10 ctgagccctc ctggcttctt catgggaccc accttagact tttaggatcc agaacaaggt 9780
 gcagggtttg cccagggcct caacatcctg tcgctgcca gctctcatat cctgctggag 9840
 accaacaagg gcccagctt cccaacagtc atggtaatcc ccagcgagat gctaaagggg 9900
 acgggagccc caggggcccg tgggcttact ggggctggtg tctcccccaca ggtgatacag 9960
 atgtacaagg aaagcccaga gatcctgagt ccaccagctc cacagggagg cgggcttagg 10020
 tggggggcat cctcctcctt gcccagagc tttgccagag tgactacctc catgaccgtg 10080
 15 gccatccacc cacccccctt ccctggtgcc tacatccatc cagcccctga tgagccccc 10140
 tggctcccctg ctgtcactag ctctggcaac ctgagttcca ggggaccagg tccagccact 10200
 gggtgaaaga gcagctgaag cacagagacc atgtgtgggg cgtgtgggg cactgggaag 10260
 cactgggtct ggtgttagac gcaggatgga cccctggagg gctctgctgc tgctgcatcc 10320
 cctctcccga cccagctgtc atgggectcc ctgatatcca tacagaacag ccaccgattt 10380
 20 gcacatccag gcctgtgtga gcctgtatct gtgtcacttg agagtgaag ctggcacttg 10440
 gggctgcagt gcagccctgt ccccttccc accccacacc actgctgtcc cagctgacca 10500
 agcctgaggg accctccagc acccttccgt ctggtgactc ctgggcaggc tctccatctc 10560
 cctgcccacc tccctaccaca tccattatct atatgaaaat gtctatcttt gtagtatata 10620
 25 tacatgttag ctatgatgaa agttttatct tttaaagaat gaaatatatt ctatgtgaag 10680
 ctatgatgaa agttttatct tttaaagaat gaaatatatt ctatgtgaag taatctcgaa 10740
 agttttatct tttaaagaat gaaatatatt ctatgtgtgc aagtgaacat tagcttcagt 10800
 tgcttttttt tggacagagt ggggagtttg caagtgaaca ttagctattg gaaggagctt 10860
 ctctggtgcc aggcctgag gtattagctt ctctagttct ggggtgaaaa gaccccgat 10920
 30 tctgtatctt tgtcatatac ttgtaacat catctggatt aagtgttac tatacaaaac 10980
 gataacaaat tttgttggtg tgaaatccta ctgggttcaa tctggagacc gagagcagaa 11040
 aaaaaagaac cccactgtgt ggctttcaga gccaccatat tccagcctgc ccgtctctcc 11100
 agactcacct ccacctacct gcttcacccg cacgggaaac ggcaaggcag aggggcaag 11160
 ccatgcagca ggtggaaggc gaggtggagg cagatcagga aagcagccag ttgaagcaga 11220
 35 gagaggctaa cagggtcttg ggagcttctc agggagtttg tggaccagag gaaaggagcc 11280
 aggttccaga gcaacctcca aggcaaggc ctctgtaagt tggttgtcct gacagccgag 11340
 aggtgtcttt ggccagtcag ccagtggatc agttgcggga actgtctcaga aactgaggtg 11400
 ctagcagtta gtgaggacac agcgtgaagt gtttgttctg tgaaagttga acagctccac 11460
 taagcagagg ccttgaagag tggccacagc cctggaatag agcacagagc ctccctaga 11520
 40 ggcgtgggga ggtttgcaac tgccccttcc cagccatagc ttaggaccca tagtctagtt 11580
 cacatagacc ctgggctcca accaccact caccaggaat gatccacccc caggaaacat 11640
 gcgttctcac atcccacccc acctggacaa aggcaggaa atcatgttct gaccaaaaga 11700
 tacaacaaca aaaaacaaca caacaaaaaa cgccatttgc aattgaatcc acgctaaaat 11760
 gcctaaaaag ctcaagagaa gcgggtagtt ggcagagaa cttagagtag ggggtgcaac 11820
 45 agcaggccca agggaggag gctgcatttg ggtccagcag tgtttgggtc accaagaagg 11880
 gccttcttag tggagcagag agagctcacc aggcagaaat agtgcaagg gggtcagccc 11940
 tcagtgccac ttaccagcgg agtaaccctg ggcaagttag ccagcctcac taagcctccc 12000
 catcttcac tttccaggcc cgaggagatc tagcctctgc ctcccacccc agcaccctcc 12060
 catcagacac aaggagcgcc actgtctgga caggctgaat tggctctcgg gtccctaatt 12120
 50 tctcatacgc cattccctct gcttagaaca ctttctcacc tccccttgat gtgaccccat 12180
 atcaccttc gaggtgaatt ggatcggatg ccatctctc caggaggggt ggggtcgtgc 12240
 ctctgtgag gtcccagtc cctgagtggt ctgtgcccgt ctgtttcccc gtccctctct 12300
 ctaagcccgg aggccttact cgggtaaggga cggcgggaca ggaccttaac ccctgggacg 12360
 aacaccagct ccgcaaggga ctccgcaccc ggcgcggccc acgggggtgcg ggtcccagga 12420
 55 ggaccagcag agaggagcat agggagcaa agggatcag tgacccatgg cttccccggg 12480
 ggcgcggaac agcccggagc gcctgtgat ttgcataccc atggtgcacc acgaaaagat 12540
 accctcaaga tgcctgcact ccctctgtgc gcgcatctt gcactgtttt agagcatgat 12600
 gcctcttaca cgcattctgt tgcataaact acataaggg agtgcgtaac acgcagcgt 12660
 ccaacaacca taagtgtgtt aagtgttagt tctccctgcg aggttcgaag cggaaagtcac 12720
 60 gaatatactc gggtttctct tcaaagcgca taaatcttcc gccttttact aaagatttcc 12780
 gtggagagaa agttgtgagt ttttattcaa ttttttgagg cctcttattt cctgaggcta 12840
 catttttaag tattaaaagt taggcaacta caaaaaaaa aaaaaa 12886