



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 357 603**

51 Int. Cl.:  
**C12N 9/64** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04775534 .3**

96 Fecha de presentación : **12.10.2004**

97 Número de publicación de la solicitud: **1678297**

97 Fecha de publicación de la solicitud: **12.07.2006**

54 Título: **Método para producir proteínas gamma-carboxiladas.**

30 Prioridad: **14.10.2003 GB 0324044**

45 Fecha de publicación de la mención BOPI:  
**28.04.2011**

45 Fecha de la publicación del folleto de la patente:  
**28.04.2011**

73 Titular/es: **ASTRAZENECA AB.**  
**151 85 Södertälje, SE**  
**MEDIMMUNE LIMITED**

72 Inventor/es: **Fenge, Christel;**  
**Lövgren, Ann y**  
**Theelin, Anders**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 357 603 T3

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

## DESCRIPCIÓN

**Campo técnico**

La presente invención se refiere a una célula huésped que comprende un vector de expresión que comprende una molécula de ácido nucleico que codifica una proteína que requiere gamma-carboxilación y secuencias control asociadas a la expresión que comprenden un primer promotor y una molécula de ácido nucleico que codifica una  $\gamma$ -glutamyl carboxilasa y secuencias control asociadas a la expresión que comprenden un segundo promotor. La invención además se refiere a un método para producir una proteína que requiere gamma-carboxilación con altos rendimientos.

**Antecedentes de la invención**

El sangrado es un problema clínico común. Se da como consecuencia de enfermedad, traumatismo, cirugía y tratamiento medicinal. Es urgente parar mecánicamente el sangrado. Esto puede ser difícil o incluso imposible debido a la localización del sangrado o porque difunda desde muchos vasos (pequeños). Los pacientes que están sangrando pueden requerir además un tratamiento con agentes que favorezcan la hemostasis. Pueden ser productos derivados de la sangre (hemoterapia), agentes que causen la liberación de agentes hemostáticos endógenos, factores de coagulación recombinantes (F), o agentes que retrasen la disolución de los coágulos de sangre.

La primera línea de tratamiento entre los productos derivados de la sangre, a menudo obtenidos del hospital local, son sangre entera para la sustitución de volumen y apoyo de la hemostasis, glóbulos rojos empaquetados para la mejora de la capacidad del transporte de oxígeno, concentrados de plaquetas para hacer crecer el número de plaquetas (si es bajo o deficiente) y plasma congelado fresco para facilitar la hemostasis (coagulación de la sangre y agregación de plaquetas). La segunda línea de productos derivados del plasma que facilitan la hemostasis son crioprecipitados de plasma, concentrados de complejo de protrombina, concentrados de complejo de protrombina activada y factores de coagulación purificados. Hoy en día están disponibles diversos factores de coagulación como proteínas recombinantes humanas, inactivos (factores de coagulación VIII y IX) y activados (factor de coagulación VIIa).

La hemofilia es un trastorno del sangrado hereditario o adquirido tanto con el factor de coagulación anormal o deficiente o anticuerpos dirigidos hacia un factor de coagulación que inhibe la función procoagulante. Las hemofilias más comunes son la hemofilia A (falta del factor de coagulación VIII) y la hemofilia B (factor IX). Los factores de coagulación simple purificados o recombinantes son el principal tratamiento de los pacientes con hemofilia. Los pacientes con anticuerpos inhibitorios presentan un problema con el tratamiento ya que también pueden neutralizar el factor de coagulación que se administra al paciente. La forma activa de la proteína C (APC) es un inhibidor de la coagulación del plasma por degradación de los factores de coagulación Va y VIIIa activados. Se ha mostrado que la APC recombinante es un tratamiento eficaz de una indebida coagulación del plasma en pacientes con sépsis.

Los factores de coagulación para uso terapéutico se pueden obtener de plasma humano, aunque el procedimiento de purificación no es simple y requiere de muchas etapas cuyos fines son, algunos, eliminar los virus contaminantes. Pero incluso con medidas de seguridad extensivas y la evaluación de productos derivados de la sangre, la contaminación con virus infecciosos o priones no puede descartarse. Debido a este riesgo, es muy deseable producir proteínas terapéuticas humanas a partir de células recombinantes cultivadas en medio sin componentes derivados de animales. Esto no siempre es inmediato ya que muchas proteínas requieren un huésped mamífero para producirse en una forma totalmente funcional, es decir para modificarse correctamente post-translacionalmente. Entre los factores de coagulación producidos comercialmente en células recombinantes están FVII (NovoSeven), FVIII (Kogenate, Recombinate, Refacto) y FIX (BeneFix) (Roddie y Ludlam. *Blood Rev.* **11**:169 - 177, 1997) y Proteína C Activa (Xigris). Uno de los principales obstáculos en la obtención de grandes cantidades de factores de coagulación humanos recombinantes completamente funcionales se basa en el dominio Gla presente en FII, FVII, FIX, FX y proteína C. Este dominio contiene residuos de ácido glutámico que están modificados post-translacionalmente por la adición de grupos carboxilo. La producción de estos factores está dificultada por el hecho de que la sobre-expresión de estos da como resultado una proteína infra-carboxilada y, por lo tanto, inactiva. Las modificaciones Gla son el resultado de la acción de una enzima dependiente de la vitamina K denominada  $\gamma$ -glutamyl carboxilasa (GGCX). Esta enzima ha sido estudiada extensivamente por muchos científicos, particularmente aquellos implicados en la investigación del factor de coagulación (documento WO-A-8803926; Wu *et al.* *Science* **254**(5038): 1634-1636, 1991; Rehemtulla *et al.*, *Proc Natl Acad Sci USA* **90**:4611-4615, 1993; Stanley *J. Biol. Chem.* **274**(24):16940-16944, 1999; Vo *et al.*, *FEBS letters* **445**:256-260, 1999; Begley *et al.*, *The Journal of Biological Chemistry* **275**(46):36245-36249, 2000; Walker *et al.*, *The Journal of Biological Chemistry* **276**(11):7769-7774, 2001; Bandyopadhyay, *et al.* *Proc Natl Acad Sci USA* **99**(3):1264-1269, 2002; Czerwiec *et al.*, *Eur J Biochem* **269**:6162-6172, 2002; Hallgren *et al.*, *Biochemistry* **41**(50):15045-15055, 2002; Harvey *et al.*, *The Journal of Biological Chemistry* **278**(10):8363-8369, 2003). Al menos dos grupos científicos han intentado enfoques para incrementar los rendimientos expresando conjuntamente GGCX con el factor de coagulación FIX, pero no tuvieron éxito (Rehemtulla, *et al.* 1993, *ibid*; Hallgren *et al.* 2002, *ibid*). Considerando el gran interés en las proteínas  $\gamma$ -carboxiladas, se puede asumir que han fallado muchos más ensayos de expresión conjunta y que por eso no han sido publicados.

Para el FII humano (protrombina) al menos 8 de cada 10 residuos Glu tienen que modificarse correctamente para obtener completamente la protrombina funcional (Malhotra, *et al.*, *J. Biol. Chem.* **260**:279-287, 1985; Seegers y Walz 'Prothrombin and other vitamin K proteins', CRC Press, 1986). Se han realizado grandes esfuerzos en la obtención de unos altos niveles de producción de rhFII usando diversos sistemas diferentes tales como células CHO, células BHK, células 293 y sistemas de expresión de virus de vacuna, pero han fallado todos o han dado como resultado un producto infra-carboxilado y por eso una protrombina funcionalmente inactiva (Jørgensen *et al.*, *J. Biol. Chem.* **262**:6729-6734, 1987; Russo *et al.*, *Biotechnol Appl Biochem* **14**(2):222-233, 1991; Fischer *et al.*, *J Biotechnol* **38**(2):129-136, 1995; Herlitschka *et al. Protein Expr. Purif.* **8**(3):358-364, 1996; Russo *et al.*, *Protein Expr. Purif.* **10**:214-225, 1997; Vo *et al.* 1999, *ibid*; Wu y Suttie *Thromb Res* **96**(2):91-98, 1999). Las primeras productividades publicadas para la protrombina recombinante carboxilada humana son bajas; 20 mg/L para protrombina mutante (Côte *et al.*, *J. Biol. Chem* **269**:11374-11380, 1994), 0,55 mg/L para protrombina humana expresada en células CHO (completamente carboxilada, Jørgensen *et al.* 1987, *ibid*), 25 mg/L en células CHO (grado de carboxilación no mostrado, Russo *et al.* 1997, *ibid*).

El documento WO 92/19636 describe la clonación e identificación de la secuencia de una carboxilasa dependiente de la vitamina K humana y bovina. La solicitud sugiere la expresión conjunta de la carboxilasa dependiente de la vitamina K y una proteína dependiente de la vitamina K en una célula huésped adecuada para preparar la proteína dependiente de la vitamina K. No se ilustra la expresión conjunta de la carboxilasa y la proteína dependiente de la vitamina K.

Son todavía necesarios mejores métodos para producir factores de coagulación de la sangre activados con altos rendimientos. La presente invención se establece para abordar esta necesidad.

### **Compendio de la invención**

De acuerdo con un primer aspecto de la invención se proporciona una célula huésped eucariota que comprende un vector de expresión que comprende una molécula de ácido nucleico que codifica una proteína que requiere gamma-carboxilación y secuencias control asociadas a la expresión que comprenden un primer promotor y una molécula de ácido nucleico que codifica una  $\gamma$ -glutamil carboxilasa y secuencias control asociadas a la expresión que comprenden un segundo promotor, en el que el primer promotor se selecciona de promotor temprano inmediato del citomegalovirus humano (hCMV), pEF-1alpha, pRSV o pUbc, y el segundo promotor se selecciona de promotor temprano inmediato de SV40, promotor FIX minimizado o promotor de quinasa timidina de HSV.

También se describe una célula que se modifica para expresar (i) una proteína que requiere gamma-carboxilación, y (ii) una  $\gamma$ -glutamil carboxilasa, en la que las proteínas (i) y (ii) se expresan en una proporción entre 10:1 y 500:1.

También se describe una célula huésped eucariota genéticamente modificada que comprende:

(i) una secuencia de polinucleótidos que codifica la proteína  $\gamma$ -glutamil carboxilasa en la que dicha proteína  $\gamma$ -glutamil carboxilasa que codifica la secuencia se une operativamente a secuencias control de la expresión que permiten la expresión de la proteína  $\gamma$ -glutamil carboxilasa por dicha célula; y

(ii) un polinucleótido que codifica una proteína que requiere carboxilación por la proteína  $\gamma$ -glutamil carboxilasa operativamente unida a secuencias control de la expresión que permiten la expresión de dicha proteína que requiere carboxilación por dicha célula;

en la que la célula es capaz de expresar la proteína  $\gamma$ -glutamil carboxilasa y la proteína que requiere carboxilación en la proporción de al menos 1:10.

El documento WO 02/29045 describe una célula huésped que comprende un vector de expresión que comprende una molécula de ácido nucleico que codifica una proteína que requiere una gamma-carboxilación y secuencias control asociadas a la expresión y, una molécula de ácido nucleico que codifica una gamma-glutamil carboxilasa y secuencias control asociadas a la expresión.

De acuerdo con un aspecto más de la invención se proporciona un vector que comprende una molécula de ácido nucleico que codifica una proteína que requiere gamma-carboxilación y secuencias control asociadas a la expresión que comprenden un primer promotor y una molécula de ácido nucleico que codifica una  $\gamma$ -glutamil carboxilasa y secuencias control asociadas a la expresión que comprenden un segundo promotor, en el que el primer promotor se selecciona de promotor temprano inmediato del citomegalovirus humano (hCMV), pEF-1alpha, pRSV o pUbc, y el segundo promotor se selecciona de promotor temprano inmediato de SV40, promotor FIX minimizado o promotor de quinasa timidina de HSV.

De acuerdo con otro aspecto más de la invención se proporciona un método para producir proteína gamma-carboxilada que comprende: (i) cultivar una célula eucariota adaptada para expresar una proteína que requiere gamma-carboxilación y  $\gamma$ -glutamil carboxilasa en una proporción de al menos 10:1, bajo condiciones adecuadas para la expresión de ambas proteínas, y (ii) aislar la proteína gamma-carboxilada. En una realización el método se usa para producir el Factor IX humano gamma-carboxilado y en otra realización el método se usa para producir

protrombina humana gamma-carboxilada. En otra realización, la proteína gamma-carboxilada producida es el Factor X humano gamma-carboxilado.

5 También se describe un método para la producción de una proteína gamma-carboxilada en una línea celular de un mamífero, que comprende la etapa de expresar, conjuntamente con dicha proteína que requiere gamma-carboxilación en la línea celular de un mamífero, una  $\gamma$ -glutamyl carboxilasa, en la que la cantidad de proteína expresada que requiere gamma-carboxilación es al menos 10 veces mayor que la cantidad de  $\gamma$ -glutamyl carboxilasa expresada, y (ii) aislar la proteína gamma-carboxilada. En una realización el método se usa para producir el Factor IX humano gamma-carboxilado y en otra realización el método se usa para producir protrombina humana gamma-carboxilada. En otra realización, la proteína gamma-carboxilada producida es el Factor X humano gamma-carboxilado.

10 También se describe una proteína gamma-carboxilada aislada producida de acuerdo con los métodos anteriores, y el uso de proteína gamma-carboxilada aislada producida de acuerdo con los métodos anteriores en terapia de coagulación o el uso de proteína gamma-carboxilada aislada producida de acuerdo con los métodos anteriores para la fabricación de un medicamento para uso en terapia de coagulación.

15 También se describe una proteína gamma-carboxilada aislada producida de acuerdo con los métodos anteriores, y el uso de proteína gamma-carboxilada aislada producida de acuerdo con los métodos anteriores en terapia de coagulación o el uso de la proteína gamma-carboxilada aislada producida de acuerdo con los métodos anteriores para la fabricación de un medicamento para uso en terapia de coagulación.

20 De acuerdo todavía con un aspecto más de la invención se proporciona un método para producir una composición farmacéutica adecuada para inducir la coagulación de la sangre o promocionar un incremento o decrecimiento de la coagulación, que comprende purificar la proteína carboxilada activa expresada desde una célula huésped adaptada para expresar una proteína que requiere gamma-carboxilación y  $\gamma$ -glutamyl carboxilasa en una proporción de entre 10:1 y 500:1 y mezclar la proteína carboxilada purificada con uno o varios vehículos farmacéuticamente aceptables o excipientes y una composición farmacéutica obtenible del método. En una realización, la proteína carboxilada activa es el Factor IX humano gamma-carboxilado y en otra realización la proteína carboxilada activa es la protrombina humana gamma-carboxilada. En otra realización, la proteína carboxilada activa es el Factor X gamma-carboxilado.

#### **Breve descripción de las Figuras**

30 La Figura 1a representa un mapa de plásmido del vector de expresión conjunta de PN32 (protrombina + GGCX).

La Figura 1b representa un mapa de plásmido del vector de expresión Ptext5 (protrombina).

La Figura 2 representa un mapa de plásmido del vector de expresión conjunta de PP6 (protrombina + GGCX).

La Figura 3a representa un mapa de plásmido del vector de expresión conjunta de F9NopA (factor IX + GGCX).

La Figura 3b representa un mapa de plásmido del vector de expresión conjunta de F9hglx (factor IX + GGCX).

#### **35 Descripción detallada de la invención**

Los inventores han analizado un diferente enfoque para la expresión de factores de coagulación dependientes de la vitamina K recombinantes apropiadamente carboxilados a altos niveles, que implica la expresión conjunta del factor de coagulación dependiente de la vitamina K y una  $\gamma$ -glutamyl carboxilasa (GGCX) en una proporción diferencial. Como ejemplo, los inventores han expresado una protrombina humana (rhFII) y GGCX humano. En vez de usar promotores fuertes para ambos rhFII y GGCX como han tratado otros (Rehemtulla *et al.*, 1993, *ibid*; Hallgren *et al.*, 2002, *ibid*), los inventores han usado una estrategia que se centra en la fuerte expresión de FII en combinación con la expresión débil o muy débil de GGCX, tal que la cantidad de GGCX expresada fue menor que 1/10 de la rhFII expresada. Para sorpresa de los inventores, esta estrategia condujo a altos niveles de rhFII secretado modificado correctamente y una buena viabilidad de las células huésped, incluso cuando las células se cultivaron en medio químicamente definido sin componente animal. Los inventores han clonado GGCX y protrombina humana en un vector de expresión de tal forma que el nivel de mRNA de protrombina excede el del mRNA de GGCX por un factor de al menos 10. Esto da como resultado una producción con un gran exceso de proteína protrombina comparado con la proteína GGCX.

50 Como un ejemplo más, los inventores han expresado rhFIX usando los mismos vectores de expresión conjunta de GGCX. Esto dio como resultado líneas celulares productoras del mRNA del factor IX a niveles que excedían los niveles de mRNA de GGCX por un factor de al menos 10 en un caso. En otra línea celular, la proporción del mRNA de factor IX: GGCX fue aproximadamente 4-5:1. Sólo la línea celular que proporcionaba una proporción de al menos 10:1 mostraba una productividad de rhFIX sustancialmente incrementada (Tabla 1).

Tabla 1. Resumen de la productividad y la proporción de mRNA proteína carboxilada:GGCX

NOMBRE DEL CLON/CONSTRUCTO	PROTEÍNA PRODUCIDA	PRODUCCIÓN * DE PROTEÍNA TOTALMENTE ACTIVA (mg/l)	PROPORCIÓN DE mRNA DE PROTEÍNA CARBOXILADA: GGCX, APROX.	FUENTE DE DATOS
P1E2/PN32	Protrombina humana	40	250:1	Ejemplo 3
B2F4/PP6		26	50:1	Ejemplo 5
H3B10/PP6		30	30:1	Ejemplo 5
E1A9/PText5		3,5	No GGCX	Ejemplo 3
N4D5/F9NopA	FIX humano	7,3	45:1	Ejemplo 7
P1G9/F9hglx		1,3	4:1	Ejemplo 7
IC4		1 <sup>α</sup>	Sin GGCX	Rehemtulla 1993, documento de EE.UU. N°. 5.460.950

\* La productividad se midió en cultivos de centrifugación bajo condiciones de cultivo similares.

<sup>α</sup> Datos de Rehemtulla 1993 y patente de EE.UU. N°. 5.460.950.

5 Se puede esperar que los factores de coagulación dependientes de la vitamina K (FII, FVII, FIX, FX y sus formas activadas FIIa o trombina, FVIIa, FIXa, FXa) producidos por el presente método de expresión conjunta con GGCX sean útiles en la prevención y el tratamiento del sangrado después de traumatismo, cirugía o enfermedades hepáticas, nefrológicas, de factores de coagulación de plaquetas o sangre (hemofilia). De la misma forma, se puede esperar que sean útiles el factor de coagulación de la proteína C y su forma activada APC en la prevención y el tratamiento de trastornos de coagulación incrementada con o sin niveles decrecientes de proteína C. El método también es aplicable a otras proteínas que requieren carboxilación post-translacional.

10 De acuerdo con un primer aspecto de la invención se proporciona una célula huésped eucariota que comprende un vector de expresión que comprende una molécula de ácido nucleico que codifica una proteína que requiere gamma-carboxilación y secuencias control asociadas a la expresión que comprenden un primer promotor y una molécula de ácido nucleico que codifica una  $\gamma$ -glutamil carboxilasa y secuencias control asociadas a la expresión que comprenden un segundo promotor, en el que el primer promotor se selecciona de promotor temprano inmediato del citomegalovirus humano (hCMV), pEF-1 $\alpha$ , pRSV o pUbc, y el segundo promotor se selecciona de promotor temprano inmediato de SV40, promotor FIX minimizado o promotor de quinasimidina de HSV.

15 En una realización, la proporción de las proteínas expresadas está entre 10:1 y 1000:1, más preferiblemente entre 10:1 y 500:1 y todavía más preferiblemente entre 25:1 y 250:1. Una proporción particularmente adecuada es aproximadamente 200:1.

20 En realizaciones separadas, la proporción de las dos proteínas expresadas puede ser al menos 10:1, 30:1, 45:1, 50:1, 100:1, 200:1, 250:1, 300:1, 400:1, 500:1 y 1000:1.

25 En una realización particular, tanto la molécula de ácido nucleico que codifica la proteína que requiere gamma-carboxilación como las secuencias control asociadas a la expresión, y la molécula de ácido nucleico que codifica la  $\gamma$ -glutamil carboxilasa y las secuencias control asociadas a la expresión se localizan sobre el mismo vector de expresión. En otra realización, estas dos moléculas de ácido nucleico se localizan en vectores de expresión separados.

De acuerdo con un aspecto más de la invención, se proporciona un ácido nucleico de acuerdo con SEQ ID NO: 14 y SEQ ID NO: 15.

30 De acuerdo con otro aspecto de la invención, se proporcionan células huésped transfectadas o transformadas con un vector que comprende la secuencia de SEQ. ID NO: 14 o SEQ ID NO: 15 para la expresión del Factor IX humano.

De acuerdo con otro aspecto de la invención, se proporciona una célula huésped capaz de expresar el factor de coagulación IX humano y enzimas gamma carboxilasas humanas, en la que el ácido nucleico que codifica

el factor de coagulación IX humano y el ácido nucleico que codifica la gamma carboxilasa se unen operativamente a secuencias control que son capaces de expresar las dos proteínas en una proporción de al menos 10:1, respectivamente.

5 También se describe una célula huésped eucariota no humana adaptada para expresar el factor de coagulación IX humano y enzimas gamma carboxilasa humanas en una proporción de al menos 10:1. El ácido nucleico que codifica el factor de coagulación IX humano y el ácido nucleico que codifica la gamma carboxilasa se unen operativamente a secuencias control que son capaces de expresar las dos proteínas en una proporción de al menos 10:1, respectivamente.

10 De acuerdo con otro aspecto de la invención, se proporciona una célula huésped que alberga ácido nucleico exógeno que comprende el factor de coagulación IX humano que codifica el ácido nucleico bajo el control del promotor hCMV y el ácido nucleico que codifica la carboxilasa humana bajo el control del promotor SV40.

También se describe un ácido nucleico de acuerdo con SEQ. ID NO: 1, SEQ. ID NO: 2 o SEQ ID NO: 3

También se describen células huésped transfectadas o transformadas con un vector que comprende la secuencia de SEQ. ID NO: 1, SEQ. ID NO: 2 o SEQ ID NO: 3 para la expresión de protrombina humana.

15 También se describen células huésped capaces de expresar protrombina humana y enzimas gamma carboxilasa humanas, en las que el ácido nucleico que codifica la protrombina humana y el ácido nucleico que codifica la gamma carboxilasa se unen operativamente a secuencias control que son capaces de expresar las dos proteínas en una proporción de al menos 10:1, respectivamente.

20 También se describe una célula huésped eucariota no humana adaptada para expresar protrombina humana y enzimas gamma carboxilasa humanas en una proporción de al menos 10:1. En una realización particular, el ácido nucleico que codifica la protrombina humana y el ácido nucleico que codifica la gamma carboxilasa se unen operativamente a secuencias control que son capaces de expresar las dos proteínas en una proporción de al menos 10:1, respectivamente.

25 De acuerdo con un aspecto más de la invención se proporciona una célula huésped que alberga ácido nucleico exógeno que comprende el ácido nucleico que codifica la protrombina humana bajo el control del promotor hCMV y el ácido nucleico que codifica la carboxilasa humana bajo el control del promotor SV40.

30 La invención se ha ilustrado usando protrombina y el factor de coagulación IX como proteínas que requieren carboxilación. Sin embargo, diversas proteínas diferentes a la protrombina y el factor IX son dependientes de una  $\gamma$ -carboxilación correcta para su completa actividad biológica. Entre las conocidas de los hombres están el factor de coagulación FVII, que en este momento solo está disponible comercialmente producido en células de mamíferos recombinantes a relativamente bajos niveles (aproximadamente 10 mg/L o menos). La presente invención será aplicable para mejorar la productividad de cualquier proteína que dependa de la  $\gamma$ -carboxilación, dichas proteínas incluyen, pero sin limitación: protrombina, factor de coagulación II (FII), factor de coagulación VII (FVII), factor de coagulación IX (FIX), factor de coagulación X (FX), proteína C, proteína S, proteína Z, proteína Bone Gla (también conocida como: BGP o osteocalcina), proteína de Matriz Gla (MGP), polipéptido 1 Gla rico en prolina (PRRG1), polipéptido 2 Gla rico en prolina (PRRG2), proteína 6 específica de detención del crecimiento (Gas 6). Otras proteínas adecuadas son: proteína tipo FXa en veneno de serpientes elápidas (subfamilia Acanthophiina) y veneno del caracol cono (*Conus textile*).

40 Cada una de estas proteínas, incluyendo sus secuencias de ácido nucleico y aminoácido, son muy conocidas. La Tabla 2 identifica secuencias representativas de formas naturales y mutantes de las diversas proteínas que se pueden usar en la presente invención.

Tabla 2.

DESCRIPCIÓN	cDNA EMBL ACC Nº.	VARIANTES DE CORTE (PROTEÍNA)	MUTACIONES	GEN EMBL ACC Nº.
Glutamato gamma carboxilasa	BC013979	2; BC013979; AF253530	1 SNP (EMBL Nº. U65896); 2 SNPs (OMIM Nº. 137167)	U65896
	V00595	1, V00595	aprox. 100 SNP's (EMBL Nº. AF478696)	AF478696
Factor VII	AF466933	4, AF466933; AF272774; AR030786; AAN60063	21 SNPs (OMIM Nº. 277500)	J02933

DESCRIPCIÓN	cDNA EMBL ACC N°.	VARIANTES DE CORTE (PROTEÍNA)	MUTACIONES	GEN EMBL ACC N°.
Factor IX	A01819	3; A01819; M19063	5 SNPs (EMBL N°. A34669; AF536327); 108 SNPs (OMIM N°. 306900)	AF536327
Factor X	BC046125	4, BC040125; AR095306; AB005892	118 SNPs (EMBL N°. M57285; AF503510); 14 SNPs (OMIM N°. 227600)	AF503510
Proteína C	BC034377	7, AB083690; AB083693; I09623; S50739; S72338	57 SNPs (EMBL N°. AF378903); 25 SNPs (OMIM N°. 176860)	AF378903
Osteocalcina	AF141310	5; AF141310; AF141310; BC033656; X04143; X51699		X04143
Proteína GLA de Matriz	BC005272	1, BC005272		
Específico detención del crecimiento 6; Factor estimulador de AXL	BC038984	1, BC038984		
Proteína Z	M55670	2; AB033749; AB033749		
Polipéptido 1 Glu rico en prolina (ácido G-carboxiglutámico)	AF009242	2; AF009242; BC030786		
Polipéptido 2 Glu rico en prolina (ácido G-carboxiglutámico)	AF009243	2; AF009243; BC026032		
Precursor de la proteína S dependiente de la Vitamina K	BC015801	1, BC015801	aprox. 100 SNPs (EMBL N°. AY308744); 8 SNPs (OMIM N°. 176880)	AY308744

5 Será apreciado que la invención no está restringida a ninguna proteína particular o proteína que codifique la secuencia de una de estas proteínas que se expresan conjuntamente. Además, y en particular con respecto a factores de coagulación de la sangre, se han descrito numerosas formas mutantes de las proteínas en la técnica. La presente invención es igualmente aplicable a estas formas mutantes, incluyendo variantes alélicas naturales, de las proteínas como es su secuencia natural. En una realización la invención se puede llevar a cabo con cualquier proteína natural o una con al menos 90%, preferiblemente al menos 95% de identidad de secuencia con ésta.

10 La identidad de secuencia entre dos secuencias se puede determinar por análisis de alineamiento informático de pares, usando programas tales como, BestFit, PILEUP, Gap o FrameAlign. La herramienta de alineamiento preferido es BestFit. En la práctica, cuando se buscan secuencias similares/idénticas a la búsqueda problema, a partir de una base de datos de la secuencia, generalmente es necesario realizar una identificación inicial de secuencias similares usando algoritmos adecuados tales como Blast, Blast2, NCBI Blast2, WashU Blast2, FastA, o Fasta3, y una matriz de evaluación tal como Blosum 62. Dichos algoritmos tratan de aproximarse muy cercanamente al algoritmo de alineamiento "oro-estándar" de Smith-Waterman. Por eso, el programa motor de

software/búsqueda preferido para uso en el ensamblaje de similaridad, es decir, cómo dos secuencias de polipéptidos primarias se alinean es Smith-Waterman. La identidad se refiere a emparejamientos directos, la similaridad proporciona sustituciones conservadoras.

5 La expresión "γ-glutamil carboxilasa" o "GGCX", según se usa en este documento, se refiere a una enzima dependiente de la vitamina K que cataliza la carboxilación de residuos de ácido glutámico.

10 Las enzimas GGCX están ampliamente distribuidas, y se han clonado a partir de muchas especies diversas tales como la ballena blanca *Delphinapterus leucas*, el pez sapo *Opsanus tau*, la gallina (*Gallus gallus*), la mixina (*Myxine glutinosa*), el límulo (*Limulus polyphemus*), y el caracol cono *Conus textile* (Begley et al., 2000, *ibid*; Bandyopadhyay et al. 2002, *ibid*). La carboxilasa del caracol cono es similar a la carboxilasa bovina y se ha expresado en células COS (Czerwiec et al. 2002, *ibid*). Otras proteínas similares a GGCX se pueden encontrar en insectos y procariontes tales como *Anopheles gambiae*, *Drosophila melanogaster* y *Leptospira* con números de entrada NCBI: gi 31217234, gi 21298685, gi 24216281, gi 24197548 y (Bandyopadhyay et al., 2002, *ibid*), respectivamente. La enzima carboxilasa muestra una remarcable conservación evolutiva. Algunas de las enzimas no humanas han mostrado, o se puede predecir que tienen, una actividad similar a la de la GGCX humana que los inventores han usado, y se pueden usar por lo tanto como una alternativa a la enzima humana.

15 La Tabla 3 identifica secuencias representativas de proteínas homólogas predichas respecto a GGXC humana (clasificadas después del origen de las especies) que se puede usar en la presente invención.

Tabla 3.

Especies	Nº ID de entrada de la base de datos
<i>Homo sapiens</i> (hombre)	NM_000821.2
	HUMGLUCARB
	HUMHGCA
	BC004422
	HSU65896
	AF253530.1
<i>Papio hamadryas</i> (pez sapo)	AC116665.1
<i>Delphinapterus leucas</i> (ballena blanca)	AF278713
<i>Bos taurus</i> (bovino)	NM_174066.2
	BOVCARBOXG
	BOVBGCA
<i>Ovis aries</i> (oveja doméstica)	AF312035
<i>Rattus norvegicus</i> (rata común)	NM_031756.1
	AF065387
<i>Mus musculus</i> (ratón)	NM_019802.1
	AF087938
<i>Opsanus tau</i> (peces óseos)	AF278714.1
<i>Conus textile</i> (moluscos)	AY0044904.1
	AF382823.2
<i>Conus imperialis</i> (moluscos)	AF448234.1
<i>Conus episcopatus</i> (moluscos)	AF448233.1
<i>Conus omaria</i> (moluscos)	AF448235.1
<i>Drosophila melanogaster</i> (mosca de la fruta)	NM_079161.2
<i>Anopheles gambiae</i> (mosquito)	XM_316389.1



Especies	Nº ID de entrada de la base de datos
<i>Secale cereale</i> (monocotiledóneas)	SCE314767
<i>Triticum aestivum</i> (trigo común)	AF280606.1
<i>Triticum urartu</i> (monocotiledóneas)	AY245579.1
<i>Hordeum vulgare</i> (cebada)	BLYHORDCA
<i>Leptospira interrogans</i> (espiroquetas)	AE011514.1
<i>Streptomyces coelicolor</i> (bacterias Gram+ alto GC)	SCO939109
	SC0939124
	AF425987.1
<i>Streptomyces lividans</i> (bacterias Gram+ alto GC)	SLU22894
<i>Streptomyces virginiae</i> ( bacterias Gram+ alto high GC)	SVSNBDE
<i>Micrococcus luteus</i> (bacterias Gram+ alto GC)	MLSPCOPER
<i>Chlamydomonas reinhardtii</i> (algas verdes)	AF479588.1
<i>Dictyostelium discoideum</i> (moho del fango)	AC115612.2
<i>Coturnix coturnix</i> (pájaros)	AF364329.1
<i>Bradyrhizobium japonicum</i> ( $\alpha$ -protoobacterias)	AP005937.1
<i>Rhodobacter sphaeroides</i> ( $\alpha$ -proteobacterias)	RSY14197
<i>Sinorhizobium meliloti</i> ( $\alpha$ -proteobacterias)	RME603647
	AF119834
<i>Mesorhizobium loti</i> ( $\alpha$ -proteobacterias)	AP003014.2
<i>Chromobacterium violaceum</i> ( $\beta$ -proteobacterias)	AE016910.1
	AE016918.1
<i>Pseudomonas aeruginosa</i> ( $\gamma$ -proteobacterias)	AE004613.1
	AF165882
<i>Xanthomonas axonopodis</i> ( $\gamma$ -proteobacterias)	AE011706.1
<i>Virus del herpes humano 8</i>	KSU52064
	KSU75698
	AF305694
	AF360120
	AF192756

Cada una de las proteínas GGCX anteriormente identificadas y proteínas GGCX de otras especies se pueden usar como enzima carboxilasa en la presente invención.

5 Una forma de realizar la expresión diferencial de las dos proteínas expresadas conjuntamente es usar diferentes promotores como parte de las secuencias control respectivas de la expresión. La técnica está repleta de ejemplos de diferentes promotores y otras secuencias control de la expresión que son capaces de expresar proteínas heterólogas en diferentes grados o dimensiones. La tecnología de expresión recombinante se lo bastante avanzada para que cualquier experto en la técnica de la expresión de proteína sea capaz de seleccionar promotores y otras secuencias control para realizar la expresión conjunta de la proteína que requiere carboxilación y la carboxilasa en la proporción deseada. La selección de dichos promotores particulares y otras secuencias control de la expresión que se usan es un asunto de elección particular.

10

En una realización, las secuencias control asociadas con la proteína que requieren la gamma-carboxilación comprenden un promotor fuerte. En una realización, es el promotor temprano inmediato del citomegalovirus humano (hCMV). Un promotor fuerte se define en este documento como un promotor que da un aumento de más de 1000 transcritos/célula. Un promotor débil se define en este documento como un promotor que da un aumento de menos de 1000 transcritos/célula.

En otra realización, las secuencias control asociadas con la  $\gamma$ -glutamyl carboxilasa comprenden un promotor débil. En una realización, el promotor temprano es el SV40. En otra realización, la proteína que requiere gamma-carboxilación y la  $\gamma$ -glutamyl carboxilasa están bajo el control de diferentes elementos promotores con el promotor que controla la expresión de la  $\gamma$ -glutamyl carboxilasa que es más débil que el promotor que controla la expresión de la proteína que requiere gamma-carboxilación.

En otra realización, la  $\gamma$ -glutamyl carboxilasa está bajo el control del promotor temprano SV40 y la proteína que requiere la gamma-carboxilación está bajo el control del promotor temprano inmediato del citomegalovirus humano (hCMV). En una realización de acuerdo con este aspecto particular de la invención, la proteína que requiere gamma-carboxilación es el Factor X humano. En otra realización, la proteína que requiere gamma-carboxilación es la protrombina humana. En otra realización, la proteína que requiere gamma-carboxilación es el Factor IX humano.

La invención se ha ilustrado mediante el uso del promotor CMV fuerte (Boshart *et al.* *Cell* **41**:521-530, 1985) para sobre-expresar el Factor IX o la protrombina y el promotor SV40 más débil (Wenger *et al.* *Anal Biochem* **221**:416-418, 1994) para controlar la expresión de GG CX. Otros promotores fuertes que se podrían usar de acuerdo con la presente invención incluyen, pero sin limitación, pEF-1 $\alpha$  [gen de la subunidad del factor-1 $\alpha$  de elongación humano] (Mizushima y Nagata, *Nuc Acids Res* **18**:5322, 1990; Goldman *et al.*, *BioTechniques* **21**:1013-1015, 1996), pRSV [virus del sarcoma de Rous (Gorman *et al.*, *Proc Natl Acad Sci USA* **79**:6777-6781, 1982)] y pUbc [ubiquitina humana (Schorpp *et al.*, *Nuc Acids Res* **24**:1787-1788, 1996)].

Es importante asegurarse de que la proteína que se produce (proteína que requiere carboxilación) lo sea en exceso comparado con la enzima de modificación, dando una proporción de al menos 10:1. Lo modos de lograr un nivel de expresión bajo de la enzima de modificación ( $\gamma$ -glutamyl carboxilasa) incluyen:

- 1) El uso de un promotor débil para controlar la expresión de la enzima de modificación incluyendo, pero sin limitación, el promotor temprano inmediato de SV40, el promotor FIX minimizado (Rouet *et al.*, *The Journal of Biological Chemistry* **267**:20765-20773, 1992) o el promotor de quinasa timidina de HSV (Wenger *et al.*, 1994, *ibid*).
- 2) Mutar las secuencias del promotor o potenciador de un promotor fuerte para reducir la fuerza del promotor.
- 3) Eliminar o cambiar la secuencia Kozak (señal de iniciación de la translación) para reducir la eficacia de la translación (Kozak. *Nuc Acids Res* **15**:8125-8148, 1987; Kozak. *Proc Natl Acad Sci USA* **87**:8301-8305, 1990).
- 4) Clonar el ácido nucleico que codifica la proteína que se produce (proteína que requiere la carboxilación) y el ácido nucleico que codifica GG CX en vectores separados y transfectar con un gran exceso del constructo que contiene la proteína que se produce de modo que se genere una célula con múltiples copias del constructo que contiene la proteína que se produce.
- 5) Clonar el DNA que codifica la proteína que se produce y el DNA que codifica los vectores de modificación de GG CX en vectores separados, transfectar conjuntamente o transfectar de forma separada, y usar un sistema de amplificación para amplificar la expresión de la proteína que se produce.
- 6) Aislar una línea celular estable que exprese recombinantemente bajos niveles de GG CX (pero por encima de los niveles endógenos) y usarlo como línea de la célula huésped para la expresión de proteínas que necesitan la  $\gamma$ -carboxilación.
- 7) Introducir mutación(ones) en la GG CX para disminuir la afinidad del sustrato de GG CX.

Además de estos, el experto en la técnica de expresión de proteínas recombinantes estará al tanto a otros métodos que se podrían usar para generar una célula huésped que exprese la proteína que requiere carboxilación y la proteína carboxilasa en una proporción de al menos 10:1.

También se describe una célula que se modifica o adapta para expresar (i) una proteína que requiere gamma-carboxilación, y (ii) una  $\gamma$ -glutamyl carboxilasa, en la que las proteínas (i) y (ii) se expresan en una proporción entre 10:1 y 500:1. En una realización particular, la  $\gamma$ -glutamyl carboxilasa se expresa entre 2 y 5 veces por encima de los niveles endógenos (es decir, los de una célula no modificada o adaptada).

También se describe una célula recombinante adaptada para expresar (i) la proteína  $\gamma$ -glutamyl carboxilasa por encima de los niveles constitutivos encontrados en una célula inadapta equivalente y (ii) una proteína que requiere carboxilación, en la que la cantidad de proteína  $\gamma$ -glutamyl carboxilasa expresada y proteína que requiere carboxilación está en la proporción de al menos 1:10.

También se describe una célula huésped eucariota genéticamente modificada que comprende:

(i) una secuencia de polinucleótidos que codifica la proteína  $\gamma$ -glutamil carboxilasa en la que dicha proteína  $\gamma$ -glutamil carboxilasa que codifica la secuencia se une operativamente a secuencias control de la expresión que permiten la expresión de la proteína  $\gamma$ -glutamil carboxilasa por dicha célula; y

5 (ii) un polinucleótido que codifica una proteína que requiere carboxilación por la proteína  $\gamma$ -glutamil carboxilasa operativamente unida a secuencias control de la expresión que permiten la expresión de dicha proteína que requiere carboxilación por dicha célula;

en la que la célula es capaz de expresar la proteína  $\gamma$ -glutamil carboxilasa y la proteína que requiere carboxilación en la proporción de al menos 1:10.

10 También se describe una célula adaptada para expresar una proteína que requiere gamma-carboxilación y  $\gamma$ -glutamil carboxilasa, en la que el ácido nucleico que codifica la proteína que requiere gamma-carboxilación y el ácido nucleico que codifica la  $\gamma$ -glutamil carboxilasa están bajo el control de secuencias reguladoras adecuadas para asegurar que la cantidad de proteína expresada que requiere gamma-carboxilación sea de al menos 10 veces la cantidad de la proteína  $\gamma$ -glutamil carboxilasa.

15 Al menos una de la proteína que requiere gamma-carboxilación y la  $\gamma$ -glutamil carboxilasa es expresada a partir del ácido nucleico que ha se ha introducido en la célula por tecnología recombinante. Un modo alternativo de trabajar la invención es expresar la proteína endógena (proteína que requiere la carboxilación o la carboxilasa), pero con sustitución de las secuencias control endógenas (promotor etc.) con secuencias heterólogas para obtener el nivel deseado de expresión.

20 La célula anfitriona es preferentemente una célula eucariota. Las células huésped típicas incluyen, pero sin limitación, células de insecto, células de levadura y células de mamífero. Las células de mamífero son particularmente preferidas. Las líneas de células de mamífero incluyen, pero sin limitación, las células CHO, HEK, NS0, 293, Per C.6, BHK y COS, y sus derivados. En una realización, la célula huésped es la de la línea celular CHO-S de mamífero.

25 La sobre-expresión de proteínas dependientes de la carboxilación ha dado como resultado recientemente de forma general productos infra-carboxilados. Esto es debido a que se está limitando la capacidad de carboxilación de la célula huésped endógena. Por otra parte, la enorme (de 16 a 70 veces) sobre-expresión de la actividad de GGCX no mejora el rendimiento del producto (Rehemtulla *et al.*, *Proc Natl Acad Sci USA* **90**: 4611-4615, 1993), (Berkner y Pudota, *Proc Natl Acad Sci USA*. **95**: 446-471, 1998), (Hallgren *et al.*, *Biochemistry* **41**(50): 15045-15055, 2002). La razón de esto no se entiende del todo. La invención de los inventores requiere una sobre-expresión moderada de GGCX. Esto asegura que se expresen de la célula niveles mayores que los endógenos de GGCX, por ejemplo, cuando el nivel de actividad de GGCX se eleve solo de 1,5 a 5 veces. En este nivel moderadamente elevado, se obtuvieron niveles sorprendentemente altos de rhFII completamente carboxilado como se muestra en el Ejemplo 1.

35 Se apreciará por lo tanto que la proporción de expresión de la proteína que requiere carboxilación y la carboxilasa que distingue esta invención de previas investigaciones de expresión conjunta, excluye los niveles de GGCX que son producidos endógenamente. Para cumplir la alta productividad requerida, es necesario expresar la carboxilasa y la proteína que requiere carboxilación a niveles por encima de los encontrados en células normales.

40 En una realización, se usa una célula o una línea celular que tiene poca o ninguna carboxilasa constitutivamente expresada y/o proteína que requiere carboxilación.

En una realización, la  $\gamma$ -glutamil carboxilasa se expresa al menos lo mismo o igual al 10% de la cantidad de la proteína que requiere gamma-carboxilación. En otras realizaciones alternativas, la  $\gamma$ -glutamil carboxilasa se expresa al menos lo mismo o igual al 5%, 2%, 1%, 0,5%, 0,25%, 0,1%, 0,05% o 0,01% de la cantidad de la proteína que requiere gamma-carboxilación.

45 El grado de expresión de las dos proteínas se puede medir usando técnicas familiares a cualquier experto en la técnica. Éstas incluyen medidas directas, por ejemplo, midiendo la actividad biológica de la proteína, o la cantidad de proteína (por ejemplo, usando anticuerpos), o medidas indirectas, por ejemplo vía la medida de niveles de transcripción del mRNA (por ejemplo, análisis Taqman como en el Ejemplo 3). Las siguientes referencias describen formas de medir la actividad de la enzima GGCX (Lingenfelter *et al.*, *Biochemistry* **35**: 8234-8243, 1996; Berkner *et al.*, *Proc Natl Acad Sci USA* **95**: 446-471, 1998; Hallgren *et al.*, *Biochemistry* **41**(50): 15045-15055, 2002; y, Berkner *et al.*, *Proc Natl Acad Sci USA* **89**: 6242-6246, 1992).

Para los fines de esta invención, la proporción de la expresión de las dos proteínas se determina indirectamente vía el nivel de transcripción del mRNA (por ejemplo, por análisis Taqman).

55 En una realización, la proteína que requiere gamma carboxilación es un factor de coagulación dependiente de la vitamina K. En otra realización más, la proteína que requiere gamma-carboxilación se selecciona preferiblemente del grupo que consiste en: protrombina, factor de coagulación II, factor de coagulación FII, factor de coagulación VII, factor de coagulación FVII, factor de coagulación IX, factor de coagulación FIX, factor de

coagulación X, factor de coagulación FX, proteína C, proteína S, proteína Z, proteína Bone Gla, proteína de Matriz Gla, proteína específica de detención del crecimiento 6 y proteína del tipo FXa de *Acanthophiinae*.

5 En una realización particular, la proteína que requiere gamma-carboxilación es el Factor IX. En otra realización particular, la proteína que requiere gamma-carboxilación es la protrombina. En otra realización, la proteína que requiere gamma-carboxilación es el Factor X.

La presente invención tiene aplicación general en proteínas que requieren la carboxilación de cualquier fuente. Sin embargo, si la proteína expresada tiene que usarse para fines terapéuticos no humanos, las proteínas humanas son particularmente preferidas.

10 En una realización, la  $\gamma$ -glutamil carboxilasa es de origen de ratón, rata, bovino o caracol cono. En otra realización, la  $\gamma$ -glutamil carboxilasa es una proteína humana.

De acuerdo con otro aspecto más de la invención, se proporciona un método para producir proteína gamma-carboxilada que comprende: (i) cultivar una célula eucariota adaptada para expresar una proteína que requiere gamma-carboxilación y  $\gamma$ -glutamil carboxilasa en una proporción de al menos 10:1, bajo condiciones adecuadas para la expresión de ambas proteínas, y (ii) aislar la proteína gamma-carboxilada.

15 También se describe un método para la producción de una proteína gamma-carboxilada en una línea celular de un mamífero, que comprende la etapa de expresar conjuntamente con dicha proteína que requiere la gamma-carboxilación en la línea celular de un mamífero una  $\gamma$ -glutamil carboxilasa, en la que la cantidad de la proteína expresada que requiere gamma-carboxilación es al menos 10 veces mayor que la cantidad de la  $\gamma$ -glutamil carboxilasa expresada; y (ii) aislar la proteína gamma-carboxilada.

20 Un método para producir una proteína gamma-carboxilada que comprende:

25 a) la modificación genética de una célula eucariota para introducir un primer polinucleótido que codifica una proteína que requiere la carboxilación y las secuencias control de la expresión que la acompañan y un segundo polinucleótido que codifica una  $\gamma$ -glutamil carboxilasa y secuencias control de la expresión que la acompañan para producir una célula huésped eucariota capaz de la expresión conjunta de la proteína que requiere carboxilación y una proteína  $\gamma$ -glutamil carboxilasa en una proporción de al menos 10:1; b) cultivar la célula en un medio de cultivo adecuado bajo condiciones que permitan que la primera y segunda secuencias de polinucleótidos sean expresadas; y c) aislar la proteína carboxilada del medio o células huésped.

30 Los vectores de expresión normalmente incluyen un origen de replicación, un promotor, un sitio de iniciación de la translación, opcionalmente un péptido señal, un sitio de poliadenilación, y un sitio de terminación de la transcripción. Estos vectores también contienen normalmente uno o más gen(es) marcador(es) de resistencia a los antibióticos para la selección. Los vectores de expresión adecuados pueden ser plásmidos, cósmidos o virus tales como fago o retrovirus. La secuencia codificante del polipéptido se sitúa bajo el control de un promotor apropiado (es decir, HSV, CMV, TK, RSV, SV40 etc), elementos de control y terminador de la transcripción (estas son las secuencias control asociadas a la expresión) de modo que la secuencia de ácido nucleico que codifica el polipéptido se transcriba en RNA en la célula huésped transformada o transfectada por el constructo del vector de expresión. La secuencia codificante puede contener o no un péptido señal o secuencia líder para la secreción del polipéptido fuera de la célula huésped. Los vectores preferidos comprenderán normalmente al menos un sitio de clonación múltiple. En ciertas realizaciones habrá un sitio de clonación o múltiples sitios de clonación situados entre el promotor y el gen que se expresa. Dichos sitios de clonación se pueden usar para crear proteínas de fusión del extremo N-terminal clonando una segunda secuencia de ácido nucleico en el sitio de clonación de modo que sea contigua y en marco con la secuencia del gen. En otras realizaciones, puede haber un sitio de clonación o múltiples sitios de clonación situados inmediatamente corriente abajo del gen para facilitar la creación de fusiones del extremo C-terminal de una forma similar a las fusiones del extremo N-terminal descritas más arriba.

45 La célula huésped se puede modificar genéticamente (tener ácidos nucleicos extra introducidos) por numerosos métodos, muy conocidos por cualquier experto en la técnica, tales como transfección, transformación y electroporación.

También se describe la proteína gamma carboxilada purificada producida por los métodos de la presente invención y su uso en terapia con coagulantes.

50 También se describe un método para promocionar el incremento o decrecimiento de la coagulación en un sujeto que comprende administrar una cantidad farmacológicamente eficaz de una proteína gamma-carboxilada aislada obtenida por los métodos descritos más arriba a un paciente que lo necesita.

55 De acuerdo con un aspecto más de la invención, se proporciona un método para producir una composición farmacéutica adecuada para inducir la coagulación de la sangre, que comprende purificar la proteína carboxilada activa expresada de una célula huésped adaptada para expresar una proteína que requiere gamma-carboxilación y  $\gamma$ -glutamil carboxilasa en una proporción de al menos 10:1 y mezclar la proteína carboxilada purificada con uno o varios vehículos farmacéuticamente aceptables o excipientes.

Los agentes terapéuticos proteicos se almacenan normalmente congelados, refrigerados, a temperatura ambiente, y/o en un estado liofilizado.

Las composiciones de la invención se pueden obtener por procedimientos convencionales usando excipientes farmacéuticos convencionales, muy conocidos en la técnica, pero lo más probablemente estarán en una forma adecuada para la inyección, tanto parenteralmente como directamente en el sitio de la herida.

Las suspensiones acuosas contienen en general el ingrediente activo en forma de polvo finamente dividido junto con uno o más agentes de suspensión, tales como carboximetilcelulosa de sodio, metilcelulosa, hidroxipropilmetilcelulosa, alginato de sodio, polivinilpirrolidona, goma de tragacanto y goma arábiga; agentes de dispersión o agentes humectantes tales como lecitina o productos de condensación de un óxido de alquileo con ácidos grasos (por ejemplo estearato de polioxietileno) o productos de condensación de óxido de etileno con alcoholes alifáticos de cadena larga, por ejemplo heptadecaetilenoxicetanol o productos de condensación de óxido de etileno con ésteres parciales procedentes de ácidos grasos y un hexitol tal como monooleato de polioxietileno-sorbitol o productos de condensación de óxido de etileno con alcoholes alifáticos de cadena larga, por ejemplo heptadecaetilenoxicetanol o productos de condensación de óxido de etileno con ésteres parciales procedentes de ácidos grasos y un hexitol tal como monooleato de polioxietileno-sorbitol o productos de condensación de óxido de etileno con ésteres parciales procedentes de ácidos grasos y anhídridos de hexitol, por ejemplo monooleato de polietileno-sorbitán. Las suspensiones acuosas también pueden contener uno o más conservantes (tales como p-hidroxibenzoato de etilo o propilo, antioxidantes (tales como ácido ascórbico), agentes colorantes, agentes saporíferos y/o agentes edulcorantes (tales como sacarosa, sacarina o aspartamo).

Se pueden formular suspensiones oleosas suspendiendo el principio activo en un aceite vegetal (tales como aceite de cacahuete, aceite de oliva, aceite de sésamo o aceite de coco) o en un aceite mineral (tal como parafina líquida). Las suspensiones oleosas también pueden contener un agente espesante tal como cera de abejas, parafina sólida o alcohol cetílico. Se pueden añadir agentes edulcorantes tales como los indicados anteriormente y agentes aromatizantes, para proporcionar una preparación oral apetitosa. Se pueden conservar estas composiciones por la adición de un antioxidante tal como ácido ascórbico.

Los polvos adecuados para la preparación de una preparación acuosa para la inyección, por la adición de un diluyente adecuado, generalmente contiene el ingrediente activo junto con vehículos adecuados y excipientes, agentes de suspensión y uno o varios estabilizadores o conservantes. El diluyente puede contener otros excipientes adecuados, tales como conservantes, modificantes de la tonicidad y estabilizantes.

Las composiciones farmacéuticas de la invención también pueden estar en forma de emulsiones de aceite en agua. La fase oleosa puede ser un aceite vegetal, tal como aceite de oliva o aceite de cacahuete o un aceite mineral, tal como por ejemplo parafina líquida o una mezcla de cualquiera de éstos. Los agentes emulsionantes adecuados pueden ser, por ejemplo, gomas que se encuentran en la naturaleza tales como goma arábiga o goma de tragacanto, fosfátidos que se encuentran en la naturaleza tales como soja, lecitina, ésteres o ésteres parciales procedentes de ácidos grasos y anhídridos de hexitol (por ejemplo monooleato de sorbitán) y productos de condensación de dichos ésteres parciales con óxido de etileno, tales como monooleato de sorbitán polioxietilénico.

Las composiciones farmacéuticas de la invención también pueden estar en la forma de una solución estéril o suspensión en un diluyente o disolvente parenteralmente aceptable no tóxico, que se puede formular de acuerdo con procedimientos conocidos usando uno o varios de los agentes de dispersión o humectación apropiados y agentes de suspensión, que se han mencionado más arriba. Una preparación inyectable, estéril, también puede ser una solución o suspensión inyectable, estéril, en un diluyente o disolvente aceptable por vía parenteral, no tóxico, por ejemplo, una disolución en 1,3-butanodiol.

Para más información sobre Formulación se remite al lector al capítulo 25.2 en el Volumen 5 de *Comprehensive Medicinal Chemistry* (Corwin Hansch; jefe del panel editorial), Pergamon Press 1990; o, Volumen 99 de *Drugs and the pharmaceutical sciences; Protein formulation and delivery* (Eugen J. McNally, editor ejecutivo), Marcel Dekker Inc 2000.

La cantidad de principio activo que se combina con uno o más excipientes para producir una forma de dosificación única variará necesariamente dependiendo del hospedador tratado y de la vía de administración particular. Por ejemplo, una formulación destinada a la inyección a humanos contendrá en general, por ejemplo, de 0,5 mg a 2 g de agente activo mezclado con una cantidad apropiada y conveniente de excipientes que puede variar desde aproximadamente 5 a aproximadamente 98 por ciento en peso de la composición total. Las formas de dosificación individuales contendrán en general de aproximadamente 1 mg a aproximadamente 500 mg del ingrediente activo. Las terapias con proteínas normalmente se liofilizan. Para más información sobre Vías de Administración y Regímenes de Dosificación se remite al lector al Capítulo 25,3 en el Volumen 5 de *Comprehensive Medicinal Chemistry* (Corwin Hansch; jefe del panel editorial), Pergamon Press 1990.

El tamaño de la dosis para propósitos terapéuticos o profilácticos de un compuesto de la invención variará naturalmente de acuerdo con la naturaleza y la gravedad de las afecciones, la edad y el sexo del animal o paciente y la vía de administración de acuerdo con principios muy conocidos de medicina. Cuando se usa un compuesto para fines terapéuticos o profilácticos, se administrará en general de manera que se reciba una dosis diaria en el intervalo, por ejemplo, de 0,5 mg a 75 mg por kg de peso corporal, dada si es necesario en dosis divididas. En

general, se administrarán dosis más bajas cuando se emplee una vía parenteral. Así, por ejemplo, para administración intravenosa, se usará en general una dosis en el intervalo de, por ejemplo, 0,5 mg a 30 mg por kg de peso corporal. De manera similar, para la administración por inhalación, se usará una dosis en el intervalo de, por ejemplo, 0,5 mg a 25 mg por kg de peso corporal.

5 La invención se ilustrará a continuación mediante los siguientes ejemplos no limitantes:

La práctica de la presente invención empleará, a menos que se indique otra cosa, métodos convencionales de biología molecular y técnicas de DNA recombinantes dentro del desarrollo de la técnica. Dichas técnicas se explican con detalle en la bibliografía. Véase, *por ejemplo*, Sambrook et al., eds., *Molecular Cloning: A Laboratory Manual* (3rd ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (2001); Ausubel et al., eds., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY (2002); Glover & Hames, eds., *DNA Cloning 3: A Practical Approach*, Vols. I, II, & III, IRL Press, Oxford (1995); Colowick & Kaplan, eds., *Methods in Enzymology*, Academic Press; Weir et al., eds., *Handbook of Experimental Immunology*, 5ª ed., Blackwell Scientific Publications, Ltd., Edinburgh, (1997).

### Ejemplo 1

#### 15 Amplificación de cDNA que codifica FII (hPT) humano y GGCX humano

Se compró mRNA de hígado humano de Clontech y se realizó la síntesis de cDNA usando el sistema Superscript de Invitrogen. El cDNA obtenido se usó como plantilla para la amplificación de FII humano usando:

cebador PTF0 5'-ATTCCTCAGTGACCCAGGAGCTGACA-3', (SEQ ID NO: 3) y

cebador PTEXT 5'-CTACTCTCCAACTGATCAATGACCTTCTGTATCCACTTCTT-3', (SEQ ID NO: 4).

20 GGCX humano se amplificó usando:

cebador hglx5, 5'-TCCGCAGAGCAATGGCGGTGTCT-3', (SEQ ID NO: 5) y

hglx3, 5'-CCAACATCTGGCCCCTTCAGAACT-3', (SEQ ID NO: 6).

El producto PCR que codifica FII se clonó directamente en el vector tratado con TA-TOPO pCDNA3.1V5/His (Invitrogen). La selección de un clon con cDNA de hFII insertado en la dirección correcta dio el constructo control de Ptext5 (Fig. 1 b). El cDNA que codifica GGCX bajo el control del promotor SV40 se obtuvo por transferencia del GGCX que codifica el fragmento de pCDNA3.1V5/His TA-TOPO al vector de pZeoSV2+ (Invitrogen), usando las enzimas de restricción *Bam*H1 y *Not*I. Los sitios de restricción *Eco*RV-*Not*I corriente abajo del inserto GGCX se eliminaron. Un fragmento de *Cl*al-*B*clI romo del plásmido pZeoSV2-GGCX resultante (que contiene el promotor SV40 y el inserto que contiene GGCX, pero no el sitio de poliadenilación y señal de poliadenilación corriente abajo de la secuencia que codifica GGCX) se clonó entonces en el sitio de restricción de *Dr*aIII romo de pCDNA3.1+ (Invitrogen). Un clon con el fragmento pSV40-GGCX insertado en tándem (igual dirección transcripcional) respecto al promotor CMV se seleccionó y un FII de *K*pnl-*Not*I romo que codifica el fragmento de Ptext5 se clonó en el sitio *Eco*RV para obtener el constructo PN32 (Fig. 1a). Las secuencias de DNA de PN32 y Ptext5 son como en el Apéndice 2. Todos los métodos de clonación estaban de acuerdo con los métodos estándar y/o los procedimientos recomendados por los fabricantes.

El constructo de PN32 contiene las siguientes características clave:

- El promotor temprano inmediato del citomegalovirus (hCMV) humano que controla la transcripción del cDNA de protrombina humano seguido de la señal de poliadenilación de la Hormona de Crecimiento Bovina (BGH) para la terminación de la transcripción eficiente y la poliadenilación de mRNA.
- 40 - El promotor temprano de SV40 que controla la transcripción del cDNA de  $\gamma$ -carboxilasa humana (GGCX) sin un sitio o señal de poliadenilación aparente.
- Otras características son las mostradas en la Figura 1 a).

Para la comparación del constructo de PText5 sin GGCX se usó (Figura 1b). La secuencia de nucleótidos de PTEXT5 se muestra en SEQ ID NO: 1. La secuencia de nucleótidos de PN32 se muestra en SEQ ID NO: 2

#### 45 Líneas celulares productoras de protrombina obtenidas

Los dos constructos en la Figura 1 se transfectaron en células CHO-S (Invitrogen). Los transfectantes estables se seleccionaron y rastrearon sin eran clones altamente productivos usando un ensayo comercialmente disponible para la actividad de protrombina (Chromogenix). En este ensayo las muestras que contienen protrombina se trataron primero con toxina de veneno de serpiente (Ecarina - disponible de Sigma) para generar trombina. La actividad de la trombina se analizó entonces por la adición de un sustrato cromogénico (S-2238) - que genera color cuando se procesa por trombina. La transfección y selección de clones se hicieron en paralelo con ambos constructos. El cultivo celular se hizo en medio DMEM que contenía 9% de suero bovino fetal inactivado por calor. Los clones obtenidos se adaptaron entonces para hacer crecer en medio sin componente animal. El clon de mejor

5 producción obtenido fue de la transfección con PN32 (FII + GGCX), que proporcionó hasta 400mg/L de protrombina recombinante humana cuando se cultivaba en medio químicamente definido sin componente animal (muy en exceso de cualquier nivel publicado). El rhFII producido recombinantemente se purificó (de acuerdo con el método descrito en Josic *et al.*, *Journal of Chromatography B*, **790**:183-197, 2003), y se fraccionó por cromatografía de intercambio iónico usando una columna de Q-Sepharose de acuerdo con técnicas estándar, para obtener rhFII completamente-carboxilado puro. Del rhFII producido del fermentador hasta 78 mg/L estaban completamente-carboxilados y tenían la misma actividad biológica que la protrombina purificada del plasma humano. La carboxilación se analizó por secuenciación del extremo N-terminal de la proteína y por ensayo de protrombinasa (Mao *et al.* *JBC*, **273**:30086-30091, 1998). La generación de trombina se activó en plasma pobre en plaquetas humano por la adición de factor de tejido, y el potencial de trombina endógeno se midió esencialmente como se describe en Stig *et al.*, (*Blood Coagulation and Fibrinolysis*, **14**:457-462, 2003).

10 El mejor clon obtenido con el constructo de PText5 dio una productividad de hasta 10 mg/L en medio químicamente definido sin componente animal, que está en el mismo intervalo publicado en la bibliografía. La acción de la protrombina completamente-carboxilada obtenida del clon de PText5 se estimó en aproximadamente el 50%. La recuperación final del rhFII completamente activo fue así al menos diez veces más alta usando el constructo de PN32 que contiene una disposición con bajo nivel de expresión de la  $\gamma$ -carboxilasa. Para cada uno de los constructos se identificaron varios clones con niveles de expresión similares.

### Ejemplo 2

#### Medición de la actividad de *ggcx* en líneas celulares de CHO

20 Dos líneas celulares CHO-S productoras de rhFII, obtenidas por transfección con el constructo PN32 (expresión conjunta de GGCX humano) y PTEXT5 (sin expresión conjunta de GGCX), respectivamente, se cultivaron en botellas de centrifugación usando medio sin proteína suplementado con 5  $\mu$ g/ml de vitamina K. Un décimo del medio de cultivo se reemplazó diariamente. Las células se recogieron después de 7 días de cultivo y se prepararon microsomas como se describe en Berkner *et al.*, (*Proc Natl Acad Sci USA* **89**: 6242-6246 1992). El FII recombinante humano se purificó a partir del sobrenadante del cultivo de las células recogidas. La actividad de GGCX se midió como se describe en Berkner y Pudota (*Proc Natl Acad Sci USA* **95**: 446-471 1998; y, Lingenfelter y Berkner (*Biochemistry* **35**: 8234-8243, 1996). Las medidas de los inventores mostraron que la actividad de GGCX era 1,5 veces mayor en la línea celular de CHO humana que expresa conjuntamente GGCX comparado con la línea celular de CHO que expresa solo rhFII, usando las mismas condiciones de cultivo.

### Ejemplo 3

#### Análisis de la expresión de mRNA por Reacción en Cadena de la Polimerasa de Transcripción Inversa a Tiempo Real (RT-PCR) de $\gamma$ -Carboxilasa y Protrombina en líneas celulares CHO-S

35 Dos líneas celulares CHO-S obtenidas por transfección estable con los constructos PN32 (FII+GGCX) y Ptext5 (solo FII) respectivamente, se cultivaron en botellas de centrifugación usando medio sin proteína suplementado con vitamina K. Las muestras de cultivo se retiraron después de 4, 5 y 6 días de cultivo para cubrir los niveles máximos estimados en la producción de mRNA. El RNA se aisló con Trizol™ de acuerdo con el protocolo suministrado por el vendedor, Invitrogen. El RNA aislado se trató con DNaseI con el kit DNA-free™ de Ambion. La síntesis de cDNA se llevó a cabo usando cebadores de hexámero aleatorios y los contenidos del kit del Sistema de Síntesis Superscript™ First-Strand para RT-PCR, Invitrogen.

40 Los cebadores y las sondas marcadas con Vic para la RT-PCR a Tiempo Real se seleccionaron usando el software Primer Express™, Applied Biosystems.

#### *Oligonucleótidos de $\gamma$ -Carboxilasa Humanos*

5'ACACCTCTGGTTCAGACCTTTCTT 3'

Cebador directo (SEQ ID NO: 7)

5' AATCGCTCATGGAAAGGAGTATTT 3'

Cebador inverso (SEQ ID NO: 8)

5' CAACAAAGGCTCCAGGAGATTGAACGC 3'

Sonda (SEQ ID NO: 9)

Longitud de amplicón 86 bp

#### *Oligonucleótidos de Protrombina Humana*

5' TGGAGGACAAAACCGAAAGAGA 3'

Cebador directo (SEQ ID NO: 10)

5' CATCCGAGCCCTCCACAA 3'

Cebador inverso (SEQ ID NO: 11)

5' CTCCTGGAATCCTACATCGACGGGC 3'

Sonda (SEQ ID NO: 12)

Longitud de amplicón 69 bp

5 Los cebadores se fabricaron en Operon/Qiagen y las sondas se pidieron a Applied Biosystems. También se usaron cebadores de control GAPDH de roedor y sondas (Applied Biosystems; Protocolo de Reactivos de Control de GAPDH TaqMan® Rodent de ABI N° 4308318)-Longitud de amplicón 177 bp. Las reacciones RT-PCR a tiempo real se realizaron en el detector de secuencia Prism™ 7700 de ABI, Applied Biosystems. La longitud esperada de los productos PCR amplificados se confirmó en geles de agarosa. Las series de dilución para investigar la eficiencia de las reacciones PCR se llevaron a cabo para todos los tres genes. Los niveles de expresión de  $\gamma$ -Carboxilasa y Protrombina se presentan respecto a la expresión del gen control, GAPDH de roedor.

**Protrombina**

	CHO-S PText5 día 4	CHO-S PText5 día 5	CHO-S PText5 día 6	CHO-S PN32 día 4	CHO-S PN32 día 5	CHO-S PN32 día 6
2 <sup>^-delta</sup> Ct	0,008014	0,076239	0,066677	0,204948	0,322343	0,364334

10

 **$\gamma$ -Carboxilasa**

	CHO-S PText5 día 4	CHO-S PText5 día 5	CHO-S PText5 día 6	CHO-S PN32 día 4	CHO-S PN32 día 5	CHO-S PN32 día 6
2 <sup>^-delta</sup> Ct	3,39E-07	0	0	0,000277	0,00159	0,001568

15

A partir de los niveles de expresión relativos, el rhFII:GGCX detectado, se calcularon proporciones de aproximadamente 74-232:1 dependiendo del día de recuperación de la muestra. Para la línea celular transfectada con PN32, la expresión conjunta de rhFII y GGCX, el número de transcrito por célula se calculó que era aproximadamente 8 para el mRNA de GGCX y aproximadamente 2000 para el mRNA de rhFII, dando así una proporción rhFII:GGCX de aproximadamente 250:1. Los transcritos de mRNA control de GAPDH/célula fueron para la misma muestra aproximadamente 4000.

**Ejemplo 4****Producción de FII humano**

20 El cDNA de FII humano y GGCX clonado en el Ejemplo 1 se insertó en pCDNA3.1 de la misma forma que en el Ej.1. Para dar los niveles de GGCX más altos, la señal de poliadenilación de pZeoSV2+ se incluyó en el fragmento de pSV40-GGCX-pA clonado en el sitio de *DraIII* romo de pCDNA3.1. Un clon con el fragmento que contiene GGCX en el orden inverso comparado con el Ej. 1 se seleccionó. La clonación del fragmento FII se hizo entonces de la misma forma que en el Ej. 1. El constructo final PP6 se muestra en la Figura 2 y la secuencia de nucleótidos PP6 se muestra en SEQ ID NO: 13.

30 Dos líneas celulares productoras de protrombina, B2F4 y H3B10 se obtuvieron transfectando CHO-S como se describe en el Ej. 1. La Protrombina de estas dos líneas celulares se purificó y caracterizó como en el Ej. 1. Los cultivos de B2F4 dieron productividades en el intervalo de 30-70 mg/L y la acción del completamente carboxilado de 55-87% (cuanto más rhFII menos completamente carboxilado). La adición de butirato dio una productividad algo más alta pero decreció la actividad de rhFII completamente carboxilado y no fue considerado beneficioso. H3B10 crece lentamente y dio una productividad de aproximadamente 50 mg/L, que era alta respecto a la cantidad de células en el cultivo, y la actividad de rhFII completamente carboxilado fue aproximadamente 60%. Comparado con la línea celular obtenida en el ejemplo 1, se produjo menos rhFII completamente carboxilado usando el constructo PP6 para una línea celular CHO. La producción de protrombina recombinante completamente activa está todavía lejos, sin embargo, por encima de anteriores niveles publicados.

**Ejemplo 5****Análisis de RT-PCR a Tiempo Real de la expresión de  $\gamma$ -Carboxilasa y Protrombina en líneas celulares de CHO-S midiendo la cantidad de mRNA**

40 Las líneas celulares B2F4 y H3B10 del ejemplo 4 se analizaron por análisis PCR a tiempo real por el mismo método y los mismos cebadores que en el Ej. 3. Las muestras de cultivo de 10 ml se recogieron en la productividad máxima para que fuera equivalente a las muestras en el Ej. 3. Para el clon H3B10 las muestras fueron desde el día 10 debido al lento crecimiento de este clon, y para el clon B2F4 las muestras fueron desde el día 6.



Tabla 4. Resultados del análisis RT-PCR a tiempo real de líneas celulares productoras de protrombina que expresan conjuntamente GGCX. Dos cultivos de centrifugación de 100 ml independientes para cada B2F4 y H3B 10 se muestrearon para el análisis RT-PCR a tiempo real.

Transcrito	Nº células total	Cantidad resultante de RNA total	RNA total usado para RT-PCR	Cantidad mRNA en RT-PCR	Número de células en RT-PCR	Ct	Copias mRNA	Copias mRNA/célula
<b>P1E2 * Día 6</b>								
PT	2,00E+07	2,39E-04	1,25E-08	2,50E-09	1,05E+03	19	2,10E+06	2005
GGCX	2,00E+07	2,39E-04	1,25E-08	2,50E-09	1,05E+03	27	8,19E+03	8
GAPDH	2,00E+07	2,39E-04	1,25E-08	2,50E-09	1,05E+03	18	4,19E+06	4010
<b>B2F4-1 Día 6</b>								
PT	1,30E+07	2,20E-04	1,25E-08	2,50E-09	7,39E+02	19,2	1,83E+06	2472
GGCX	1,30E+07	2,20E-04	1,25E-08	2,50E-09	7,39E+02	24,1	6,11E+04	83
GAPDH	1,30E+07	2,20E-04	1,25E-08	2,50E-09	7,39E+02	19,8	1,20E+06	1631
<b>B2F4-2 Día 6</b>								
PT	1,10E+07	1,40E-04	1,25E-08	2,50E-09	9,82E+02	19,2	1,83E+06	1859
GGCX	1,10E+07	1,40E-04	1,25E-08	2,50E-09	9,82E+02	24,1	6,11E+04	62
GAPDH	1,10E+07	1,40E-04	1,25E-08	2,50E-09	9,82E+02	19	2,10E+06	2135
<b>H3B10-1 Día 10</b>								
PT	1,10E+07	2,90E-04	1,25E-08	2,50E-09	4,74E+02	17,77	4,92E+06	10375
GGCX	1,10E+07	2,90E-04	1,25E-08	2,50E-09	4,74E+02	23,4	9,93E+04	210
GAPDH	1,10E+07	2,90E-04	1,25E-08	2,50E-09	4,74E+02	17,96	4,31E+06	9095
<b>H3B10-2. Día 10</b>								
PT	8,90E+06	3,10E-04	1,25E-08	2,50E-09	3,59E+02	19,2	1,83E+06	5087
GGCX	8,90E+06	3,10E-04	1,25E-08	2,50E-09	3,59E+02	25,3	2,66E+04	74
GAPDH	8,90E+06	3,10E-04	1,25E-08	2,50E-09	3,59E+02	18,9	2,25E+06	6263
* Datos de P1E2 del ejemplo 3 para comparación.								

La proporción calculada de mRNA de rhFII: El mRNA de GGCX fue aproximadamente 30:1 para el clon H3B10, aproximadamente 50:1 para el clon B2F4 y aproximadamente 250:1 para el clon P1E2.

**Ejemplo 6**

Producción de factor IX de coagulación humano (FIX)

5 El cDNA del factor IX de coagulación humano se amplificó a partir del cDNA de hígado del reservorio de gen humano comprado de Invitrogen. Los cebadores de oligonucleótidos fueron para:

el extremo 5'; F9f.ampl.: 5'-CACCATGCAGCGCGTGAACATGAT-3' (SEQ ID NO: 16),

y el extremo 3'; F9r.ampl.: 5'- CCTTGGAAATCCATCTTTCATTA -3' (SEQ ID NO: 17).

10 La clonación de la secuencia correcta fue confirmada por la secuenciación del DNA. El fragmento de FIX humano fue amplificado por PCR usando la polimerasa Pfx (Invitrogen) y los cebadores de clonación para producir un fragmento del extremo romo. El fragmento del extremo romo se fosforiló usando T4 polinucleótido cinasa, y se clonó en los vectores de pCDNA-GGCX EcoRV digeridos y de-fosforilados del Ej. 1 y Ej. 4. De este modo, se obtuvieron los constructos para la expresión conjunta de análogos de FIX y GGCX humano respecto a los constructos de expresión conjunta usados para la producción de protrombina humana (Ej.1 y 4). La clonación de las secuencias correctas se confirmó por la secuenciación del DNA y la expresión transitoria en células COS-7. El constructo del vector F9NopA se puede ver en la Fig. 3a y el constructo del vector F9hglx se muestra en la Fig. 3b. La diferencia entre los vectores F9NopA y F9hglx es la dirección de transcripción del gen de GGCX. La secuencia de nucleótidos de F9NopA se muestra en SEQ ID NO: 14 y la secuencia de nucleótidos de F9hglx se muestra en SEQ ID NO: 15

20 Establecimiento de líneas celulares productoras de rhFIX

Los constructos de rhFIX se transfectaron en células CHO-S usando el procedimiento descrito en el Ej. 1. Para cada constructo de FIX se rastrearon aproximadamente 3000 clones para la expresión de rhFIX por ELISA de sobrenadantes de células. Los anticuerpos usados fueron de Haemathology Technology Inc. y DakoCytomation. Los clones se seleccionaron y se adaptaron para hacerlos crecer en medio de CHO químicamente definido sin proteína. Las células se cultivaron tanto en matraces en T a 37 °C como en botellas de centrifugación a 32-37 °C. La concentración de CO<sub>2</sub> fue 5% para ambos tipos de cultivos. El rhFIX producido se purificó hasta homogeneidad por cromatografía de intercambio aniónico de Q-Sepharose a pH 7,0. La actividad de hFIX recombinante se determinó por ensayo de coagulación usando plasma deficiente de FIX (Precision Biologic). El clon de rhFIX de mejor producción obtenido fue N4D5, que se obtuvo usando el constructo de F9NopA, produjo hasta 4 µg/ml de rhFIX activo cultivado en medio químicamente definido sin proteína en matraz T. El crecimiento en la botella de centrifugación del mismo clon produjo hasta 7,1 µg/ml de rhFIX. La productividad global, incluyendo también el rhFIX carboxilado incompletamente, no activo, se estimó por transferencia Western que era al menos 30 µg/ml. El clon de mejor producción obtenido con el constructo de rhFIX F9hglx fue P1G9 que produjo 0,7 (matraz T) - 1,3 (de centrifugación) µg/ml rhFIX bajo condiciones similares. Los resultados indican que la productividad de rhFIX mejoró por la expresión conjunta de GGCX a un bajo nivel usando el constructo de F9NopA, pero que la expresión conjunta de GGCX usando el constructo F9hglx, fue menos beneficiosa. También se observó que el constructo de F9NopA, que aumentaba el clon de N4D5, generalmente dio señales de ELISA mayores que el constructo de F9hglx, que aumentaba el clon de P1G9, en rastreos simultáneos según la productividad durante el procedimiento de desarrollo de la línea celular.

40 La productividad de la línea celular de N4D5 es aproximadamente 4-6 mejor que los niveles previamente publicados obtenidos bajo condiciones comparables, en los que IC4, IGB, r-FIX BHK y r-FIX 293 son el nombre de los clones mencionados en las referencias (Tabla 5).

Tabla 5. Comparación de la productividad a partir de líneas celulares productoras de FIX humano.

Línea celular/constructo	Cantidad de rhFIX activo producido		Productividad total (µg/ml)	Referencia
	Matraz T (µg/ml)	Centrifugado (µg/ml)		
N4D5/F9NopA CHO, bajo expr. conjunta de GGXC	4	7,1	>30	Ejemplo 6
P1G9/F9hglx CHO, medio expr. conjunta de GGCX	0,7	1,3	Nd	Ejemplo 6

<b>IC4</b> CHO, HA (control) expr. conjunta	0,9	nd	30	Rehemtulla 1993.
<b>IC4</b> CHO. Alta expr. conjunta de GGCX	1	nd	29	Rehemtulla 1993.
<b>IC4</b>	0,9	nd	20	Documento de EE.UU. N°. 5.460.950
<b>1G8</b> CHO	1,5	nd	43	Kaufman, RJ et al 1986 JBC 261:9622-9628
<b>r-FIX BHK</b>	0,004/24h	nd	0,004/24h	Hallgren 2002
<b>r-FIX 293</b>	0,004/24h	nd	0,004/24h	Hallgren 2002

### Ejemplo 7

#### Análisis de RT-PCR a Tiempo Real de la expresión de $\gamma$ -Carboxilasa y Factor IX en líneas celulares de CHO-S midiendo la cantidad de mRNA

5 Se cultivaron clones productores de hFIX recombinantes en botellas de centrifugación a 32-37 °C, en 100 ml de medio químicamente definido sin proteína suplementado con Vitamina K. Las muestras de 5-10 ml se recogieron a la concentración máxima de rhFIX y se analizaron según el contenido de transcritos de FIX humano y GGCX, así como transcritos del gen de control de GAPDH (interno). El procedimiento fue como en el ejemplo 3. Los cebadores para rhFIX fueron como sigue:

#### 10 Cebadores del Factor IX Humano

5' AATAGTGCTGATAACAAGGTGGTTTG 3'

Cebador directo (SEQ ID NO: 18)

5' CACTGCTGGTTCACAGGACTTCT 3'

Cebador inverso (SEQ ID NO: 19)

5' TCCTGTACTGAGGGATATCGACTTGCAGAAAAC 3'

Sonda (SEQ ID NO: 20)

Longitud de amplicón 84 bp

15 Se encontraron niveles de RNA Mensajero máximos en diferentes días dependiendo de la temperatura del cultivo y del tamaño del inóculo del cultivo. Los niveles máximos de mRNA se encontraron que correspondían bien con la concentración máxima de rhFIX en el medio de cultivo.

Tabla 6. Resultados del análisis RT-PCR a Tiempo Real de clones productores de rhFIX.

Lote de la línea celular	Día de cultivo	2 <sup>-delta</sup> Ct FIX	2 <sup>-delta</sup> Ct GGCX	proporción de mRNA FIX:GGCX
N4D5-100	11	0,255253	0,005461	47:1
N4D5-2	14	0,264866	0,006201	43:1
P1G9-A	6	0,022982	0,005601	4:1
P1G9-B	8	0,04181	0,007687	5:1

20 A partir de los análisis RT-PCR a Tiempo Real, los inventores también encontraron que, aunque los valores de 2<sup>-delta</sup> Ct variaban con el tiempo de cultivo y la condiciones, las proporciones de mRNA de FIX:GGCX fueron aproximadamente iguales para cada clon. Para el clon que mejor producía rhFIX N4D5 la proporción fue aproximadamente 45:1. El análisis de otro clon, P1G9, dio una proporción menor de aproximadamente 4,5:1. El clon P1G9 produjo solo 20% de la cantidad rhFIX producida por N4D5.

Listado de Secuencias:

SEQ ID NO: 1

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCAGTCTCAGTACAATCTGCTCTGAT  
 GCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTG  
 CGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATC  
 TGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGA  
 CATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCAT  
 ATATGGAGTTCGCGGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGCCCAAC  
 GACCCCGGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC  
 TTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCA  
 AGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCGCT  
 GGCATTATGCCCAGTACATGACCTTATGGGACTTTTCTACTTGGCAGTACATCTACGTAT  
 TAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGATAG  
 CGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTTGT  
 TTGGCACCAAAATCAACGGGACTTTCCAATGTGCGTAACAACCTCCGCCCATTTGACGC  
 AAATGGGCGGTAGGCGGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAAC  
 TAGAGAACCCTACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCC  
 AAGCTGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAATT  
 GCCCTTATTCTCAGTGACCCAGGAGCTGACACACTATGGCGCACGTCCGAGGCTTGC  
 AGCTGCCTGGCTGCCTGGCCCTGGCTGCCCTGTGTAGCCTTGTGCACAGCCAGCATGT  
 GTTCTTGGCTCCTCAGCAAGCACGGTTCGCTGCTCCAGCGGGTCCGGCGAGCCAACAC  
 CTTCTTGGAGGAGGTGCGCAAGGGCAACCTGGAGCGAGAGTGCCTGGAGGAGACGTG  
 CAGCTACGAGGAGGCTTCGAGGCTCTGGAGTCTCCACGGCTACGGATGTGTTCTGG  
 GCCAAGTACACAGCTTGTGAGACAGCGAGGACGCTCGAGATAAGCTTGTGCATGTC  
 TGGAAGGTAACCTGTGCTGAGGGTCTGGGTACGAACCTACCGAGGGCATGTGAACATCAC  
 CCGGTCAGGCATTGAGTGCCAGCTATGGAGGAGTCGCTACCCACATAAGCCTGAAATC  
 AACTCCAATCCCATCCTGGGGCCGACCTACAGGAGAATTTCTGCCGCAACCCCGACA  
 GCAGCACCACGGGACTCTGGTGTACTACACTACAGACCCACCGTGAGGAGGACAGGAA  
 GCAGCATCCCTGTCTGTGGCCAGGATCAAGTCACTGTAGCGATGACTCCACGCTCCGA  
 AGGCTCCAGTGTGAATCTGTACCTCCATTGGAGCAGTGTGTCCCTGATCGGGGGCAG  
 CAGTACCAGGGGCGCCTGGCGGTGACCACACATGGGCTCCCCTGCCTGGCCTGGGCC  
 AGCGCACAGGCCAAGGCCCTGAGCAAGCACCAGGACTTCAACTCAGCTGTGCAGCTGG  
 TGGAGAACTTCTGCCGCAACCCAGACGGGGATGAGGAGGGCGTGTGGTGCTATGTGG  
 CCGGGAAGCCTGGCGACTTTGGGTACTGCGACCTCAACTATTGTGAGGAGGCCGTGGA  
 GGAGGAGACAGGAGATGGGCTGGATGAGGACTCAGACAGGGCCATCGAAGGGCGTAC  
 CGCCACCAGTGAGTACCAGACTTTCTTCAATCCGAGGACCTTTGGCTCGGGAGAGGCA  
 GACTGTGGGCTGCGACCTCTGTTTCGAGAAGAAGTCGCTGGAGGACAAAACCGAAAGAG  
 AGCTCCTGGAATCCTACATCGACGGGCGCATTGTGGAGGGCTCGGATGCAGAGATCGG  
 CATGTCACCTTGGCAGGTGATGCTTTTCCGGAAGAGTCCCCAGGAGCTGCTGTGTGGG  
 GCCAGCCTCATCAGTGACCGCTGGGTCTCACCGCCGCCACTGCCTCCTGTACCCGC  
 CCTGGGACAAGAACTTCACCGAGAATGACCTTCTGGTGCGCATTGGCAAGCACTCCCG  
 CACCAGGTACGAGCGAAACATTGAAAAGATATCCATGTTGGAAAAGATCTACATCCACC  
 CCAGGTACAACCTGGCGGGAGAACCTGGACCGGGACATTGCCCTGATGAAGCTGAAGAA  
 GCCTGTTGCCTTCAGTGACTACATTCACCTGTGTGTCTGCCCGACAGGGAGACGGCA  
 GCCAGCTTGTCCAGGCTGGATAAAGGGGCGGGTGACAGGCTGGGGCAACCTGAAG  
 GAGACGTGGACAGCCAACGTTGGTAAGGGGCGAGCCAGTGTCTGCAGGTGGTGAAC  
 CTGCCATTGTGGAGCGGCCGGTCTGCAAGGACTCCACCCGGATCCGCATCACTGACA  
 ACATGTTCTGTGCTGTTACAAGCCTGATGAAGGGAAACGAGGGGATGCCTGTGAAGG  
 TGACAGTGGGGGACCCCTTTGTCATGAAGAGCCCTTTAACAAACCGCTGGTATCAAAATGG  
 GCATCGTCTCATGGGGTGAAGGCTGTGACCGGGATGGGAAATATGGCTTCTACACACA  
 TGTGTTCCGCCTGAAGAAGTGGATACAGAAGGTCATTGATCAGTTTTGGAGAGTAGAAGG  
 GCAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCCCCGCGTT

CGAAGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGCGTACCGGTCATC  
 ATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGC  
 CAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCCTTGACCCTGGAAGGTGCCACTCC  
 CACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTC  
 TATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAG  
 CAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTG  
 GGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTGT  
 GGTGGTTACGCGCAGCGTGACCGCTACACTTGGCAGCGCCCTAGCGCCCCTCCTTTTC  
 GCTTTCCTCCCTTCCCTTCTCGCCACGTTGCGCCGGCTTTCCCCGTCAAGCTCTAAATCG  
 GGGCATCCCTTAGGGTCCGATTTAGTGCTTACGGCACCTCGACCCCAAAAACTTG  
 ATTAGGTGATGCTTACGTTACGTTAGTGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTG  
 ACGTTGGAGTCCAGCTTCTTAAATAGTGGACTCTTGTTCCAAACCTGGAACAACACTCAAC  
 CCTATCTCGGTCTATTCTTTTGAATTAATAAGGGATTTTGGGGATTTCCGCCATTGGTTAA  
 AAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTA  
 GGGTGTGGAAGTCCCAGGCTCCCAGGCAGGCAGAAGTATGCAAAGCATGCATCTC  
 AATTAGTCAGCAACCAGGTGTGGAAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGC  
 AAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCG  
 CCCCTAACTCCGCCAGTTCGCCATTCTCGCCCCATGGCTGACTAATTTTTTTTATT  
 TATGCAGAGGCCGAGGCCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCT  
 TTTTGGAGGCCTAGGCTTTTGCAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGAT  
 CTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCA  
 GGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCACAACAGACAA  
 TCGGCTGCTGTAGTCCGCGGTGTTCCGGTGTGACTGCAGCGCAGGGGGCGCCGGTTCTTTT  
 TGTCAAGACCGACCTGTCCGGTGCCTGAATGAAGTGCAGGACGAGGCAGGCAGGCTA  
 TCGTGGCTGGCCACGACGGGCGTTCTTGCAGCTGTGCTCGACGTTGTCAGTGAAG  
 CGGGAAGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCA  
 CCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACG  
 CTTGATCCGGCTACCTGCCATTCCGACCACCAAGCGAAACATCGCATCGAGCGAGCAC  
 GACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGG  
 GCTCGCGCCAGCCGAACTGTTCCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGA  
 TCTCGTCTGACCCATGGCGATGCCTGCTTCCGAATATCATGGTGGAAAATGGCCGC  
 TTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAG  
 CGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCCT  
 CGTGCTTACGGTATCGCCGCTCCCGATTCCGACGCGCATCGCCTTCTATCGCCTTCTTG  
 ACGAGTCTTCTGAGCGGGACTCTGGGTTCCGCGAAATGACCGACCAAGCGACGCCCA  
 ACCTGCCATCACGAGATTTGATTCACCCGCGCCTTCTATGAAAGTTGGGCTTCGGA  
 ATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGT  
 TCTTCGCCACCCCAACTTGTATTGTCAGCTTATAATGGTTACAAATAAAGCAATAGCA  
 TCACAAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACT  
 CATCAATGTATCTTATCATGTCTGTATACCGTGCACCTCTAGCTAGAGCTTGGCGTAATC  
 ATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCACACAACATACG  
 AGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAA  
 TTGCGTTGCGCTCACTGCCCGCTTCCAGTCCGGAAACCTGTCGTGCCAGCTGCATTAA  
 TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGCTATTGGGCGCTTTCGCCCTTCCCT  
 CGCTCACTGACTCGCTGCGCTCGGTCTGCTTCCGGCTGCGGCGAGCGGTATCAGCTCACTC  
 AAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGA  
 GCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC  
 ATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCG  
 AAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGC  
 TCTCCTGTTCCGACCCTGCCGCTTACGGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG  
 CGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTCCGCT  
 CCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGCCTTATCCGG  
 TAACTATCGTCTTGTGATCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCA  
 CTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG

GTGGCCTAACTACGGCTACACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC  
 CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGT  
 AGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGA  
 AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAG  
 GGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAATG  
 AAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTA  
 ATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTC  
 CCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAA  
 TGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGC  
 CGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCAGTCTATT  
 AATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT  
 TGCCATTGCTACAGGCATCGTGGTGCACGCTCGTCTGTTTGGTATGGCTTCAATCAGCT  
 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTT  
 AGCTCCTTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTTATCACTCAT  
 GGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGT  
 GACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCT  
 CTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTC  
 ATCATTGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATC  
 CAGTTTCGATGTAACCCACTCGTGCACCCAACCTGATCTTCAGCATCTTTTACTTTACCAG  
 CGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAAATGCCGCAAAAAAGGGAATAAGGGCG  
 ACACGGAAATGTTGAATACTCATACTCTTCTTTTTCAATATTATTGAAGCATTATCAGG  
 GTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGG  
 TTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC //

SEQ ID NO: 2

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACCTCTCAGTACAATCTGCTCTGAT  
 GCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTG  
 CGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATC  
 TGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGA  
 CATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTTCATAGCCCAT  
 ATATGGAGTTCGCGGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGCCCAAC  
 GACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC  
 TTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC  
 AAGTGTATCATATGCCAAGTAGCAGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCC  
 TGGCATTATGCCAGTACATGACCTTATGGGACTTTCTACTTGGCAGTACATCAATGGGCGTGGATA  
 TTAGTCATCGCTATTACCATGGTGTATGCGTTTTGGCAGTACATCAATGGGCGTGGATA  
 GCGGTTTACTCACGGGGATTTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTTGT  
 TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACG  
 CAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAA  
 CTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACC  
 CAAGCTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTG  
 TGGTGGAATTGCCCTTATTCTCAGTGACCCAGGAGCTGACACACTATGGCGCACGTCC  
 GAGGCTTGACGCTGCCTGGCTGCCTGGCCCTGGCTGCCCTGTGTAGCCTTGTGCACAG  
 CCAGCATGTGTTCTGGCTCCTCAGCAAGCACGGTGCCTGCTCCAGCGGGTCCGGCGA  
 GCCAACACCTTCTTGAGGAGGTGCGCAAGGGCAACCTGGAGCGAGAGTGCCTGGAG  
 GAGAGTGCAGCTACGAGGAGGCCTTCGAGGCTTGGAGTCCCTCCACGGCTACGGAT  
 GTGTTCTGGGCCAAGTACACAGCTTGTGAGACAGCGAGGACGCCTCGAGATAAGCTTG  
 CTGCATGTCTGGAAGGTAAGTGTGCTGAGGGTCTGGGTACGAACTACCGAGGGCATGT  
 GAACATCACCCGGTACAGGCATTGAGTGCCAGCTATGGAGGAGTGCCTACCCACATAAG  
 CCTGAAATCAACTCCACTACCCATCCTGGGGCCGACCTACAGGAGAATTTCTGCCGCAA  
 CCCCAGACAGCAGCACCACGGGACCCTGGTGTACTACTACAGACCCACCGTGAGGAG  
 GCAGGAATGCAGCATCCCTGTCTGTGGCCAGGATCAAGTCACTGTAGCGATGACTCCA  
 CGCTCCGAAGGCTCCAGTGTGAATCTGTCACCTCCATTGGAGCAGTGTGTCCCTGATC

GGGGGCAGCAGTACCAGGGGGCGCCTGGCGGTGACCACACATGGGGCTCCCCTGCCTGG  
CCTGGGCCAGCGCACAGGCCAAGGCCCTGAGCAAGCACCAGGACTTCAACTCAGCTGT  
GCAGCTGGTGGAGAACTTCTGCCGCAACCCAGACGGGGATGAGGAGGGCGTGTGGTG  
CTATGTGGCCGGGAAGCCTGGCGACTTTGGGTAAGTGCACCTCAACTATTGTGAGGAG  
GCCGTGGAGGAGGAGACAGGAGATGGGCTGGATGAGGACTCAGACAGGGCCATCGAA  
GGGCGTACCGCCACCAGTGAGTACCAGACTTCTTCAATCCGAGGACCTTTGGCTCGG  
GAGAGGCAGACTGTGGGCTGCGACCTCTGTTTCGAGAAGAAGTCGCTGGAGGACAAAAC  
CGAAAAGAGAGCTCCTGGAATCCTACATCGACGGGCGCATTGTGGAGGGCTCGGATGCA  
GAGATCGGCATGTCACCTTGGCAGGTGATGCTTTTCCGGAAGAGTCCCAGGAGCTGC  
TGTGTGGGGCCAGCCTCATCAGTGACCGCTGGGTCTCACCGCCGCCACTGCCTCCT  
GTACCCGCCCTGGGACAAGAATTACCGAGAATGACCTTCTGGTGCGCATTGGCAAG  
CACTCCCGCACCAGGTACGAGCGAAAACATTGAAAAGATATCCATGTTGAAAAGATCTA  
CATCCACCCCAGGTACAACCTGGCGGGAGAACCCTGGACCGGGACATTGCCCTGATGAAG  
CTGAAGAAGCCTGTTGCCCTTCACTGACTACATTACCCTGTGTGTCTGCCCGACAGGGA  
GACGGCAGCCAGCTTGTCCAGGCTGGATACAAGGGGGCGGGTGACAGGCTGGGGCAA  
CCTGAAGGAGACGTGGACAGCCAACGTTGGTAAGGGGGCAGCCCAGTGTCTGCAGGT  
GGTGAACCTGCCATTGTGGAGCGGCCGGTCTGCAAGGACTCCACCCGGATCCGCATC  
ACTGACAACATGTTCTGTGCTGGTTACAAGCCTGATGAAGGGAAACGAGGGGGATGCCT  
GTGAAGGTGACAGTGGGGGACCCTTTGTTCATGAAGAGCCCTTTAACAACCGCTGGTA  
TCAAATGGGCATCGTCTCATGGGGTGAAGGCTGTGACCGGGATGGGAAATATGGCTTC  
TACACACATGTGTTCCCGCTGAAGAAGTGATACAGAAGGTCATTGATCAGTTTTGGAGA  
GTAGAAGGGCAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGC  
CCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTT  
TGCCCCTCCCCGTCCTTCTGACCCTGGAAGGTGCCACTCCCCTGTCTTTCTCTA  
ATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCACTTCTATTCTGGGGGGTG  
GGGTGGGGCAGGACAGCAAGGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGG  
GATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGG  
TATCCCACGCGCCCTGTAGCGGCGCATTAAAGCGCGCGGGTGTGGTGGTTACGCGC  
AGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTGCTTTCTTCCCTTC  
CTTTCTCGCCACGTTCCGCCGCTTTCCCGCTCAAGCTCTAAATCGGGGGCTCCCTTAG  
GGTCCGATTTAGTGCTTTACGGCACCTTCGACCCCAAAAACCTTGATTAGGGCTGTGG  
AATGTGTGTCACTTAGGGTGTGGAAAGTCCCGAGGCTCCCAGCAGGCAGAAAGTATGC  
AAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCAGGCTCCCAGCA  
GGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAAC  
TCCGCCCATCCCGCCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGA  
CTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTCCAGAA  
GTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCTCTGGCTAACTAG  
AGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAG  
CTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGT  
GGAATTGCCCTTTCCGCAGAGCAATGGCGGTGTCTGCCGGTCCGCGCGGACCTCGC  
CCAGCTCAGATAAAGTACAGAAAGACAAGGCTGAACTGATCTCAGGGCCCAGGCAGGA  
CAGCCGAATAGGGAAACTCTTGGGTTTTGAGTGGACAGATTTGTCCAGTTGGCGGAGG  
CTGGTGACCTGCTGAATCGACCAACGGACCCTGCAAGCTTAGCTGTCTTTGCTTTCT  
TTTTGGGTTCTTGATGGTGCTAGACATTCCCCAGGAGCGGGGGCTCAGCTCTCTGGAC  
CGGAAATACCTTGATGGGCTGGATGTGTGCCGCTTCCCCTTGCTGGATGCCCTACGCC  
CACTGCCACTTGACTGGATGTATCTTGTCTACACCATCATGTTTCTGGGGGCACTGGGC  
ATGATGCTGGGCCTGTGCTACCGGATAAGCTGTGTGTTATTCTGCTGCCATACTGGTA  
TGTGTTTCTCCTGGACAAGACATCATGGAACAACCACTCCTATCTGTATGGGTTGTTGG  
CCTTTCAGCTAACATTCATGGATGCAAACCACTACTGGTCTGTGGACGGTCTGCTGAAT  
GCCCATAGGAGGAATGCCACGTCGCCCTTTGGAACATATGCAGTGCTCCGTGGCCAGA  
TCTTCATTGTGTACTTCATTGCGGGTGTGAAAAGCTGGATGCAGACTGGGTTGAAGGC  
TATTCCATGGAATATTTGTCCCGGCACTGGCTCTTCACTCCCTTCAAACCTGCTGTTGTCT  
GAGGAGCTGACTAGCCTGCTGGTCTGCACTGGGGTGGGCTGCTGCTTGACCTCTCAG  
CTGGTTTCTGCTCTTTTTTATGATGTCTCAAGATCCATTGGCCTGTTCTTTGTGCTCACTT

CCACTGCATGAATCCCAGCTTTTCAGCATTGGTATGTTCTCCTACGTCATGCTGGCCA  
 GCAGCCCTCTCTTCTGCTCCCCTGAGTGGCCTCGGAAGCTGGTGTCTACTGCCCCCG  
 AAGGTTGCAACAACCTGTTGCCCTCAAGGCAGCCCTCAGCCAGTGTTCCTGTGTGT  
 ATAAGAGGAGCCGGGGCAAAGTGGCCAGAAGCCAGGGCTGCGCCATCAGCTGGGAG  
 CTGCCTTACCCTGCTCTACCTCCTGGAGCAGCTATTCCTGCCCTATTCTCATTTTCTCA  
 CCCAGGGCTATAACAACCTGGACAAATGGGCTGTATGGCTATTCTGGGACATGATGGTG  
 CACTCCCGCTCCCACCAGCACGTGAAGATCACCTACCGTGATGGCCGCACTGGCGAAC  
 TGGGCTACCTTAACCCTGGGGTATTTACACAGAGTCGGCGATGGAAGGATCATGCAGA  
 CATGCTGAAGCAATATGCCACTTGCCTGAGCCGCTGCTTCCAAGTATAATGTCACTG  
 AGCCCCAGATCTACTTTGATATTTGGGTCTCCATCAATGACCGCTTCCAGCAGAGGATTT  
 TTGACCCTCGTGTGGACATCGTGCAGGCCGCTTGGTCACCCTTTCAGCGCACATCCTG  
 GGTGCAACCCTCTTGATGGACCTGTCTCCCTGGAGGGCCAAGTTACAGGAAATCAAG  
 AGCAGCCTAGACAACCACACTGAGGTGGTCTTCATTGCAGATTTCCCTGGACTGCACTT  
 GGAGAATTTTGTGAGTGAAGACCTGGGCAACACTAGCATCCAGCTGCTGCAGGGGGAA  
 GTGACTGTGGAGCTTGTGGCAGAACAGAAGAACCAGACTCTTCGAGAGGGGAGAAAAA  
 TGCAGTTGCCTGCTGGTGAAGTACCATAAGGTGTATACGACATCACCTAGCCCTTCTTGC  
 TACATGTACGTCTATGTCAACACTACAGAGCTTGCCTGGAGCAAGACCTGGCATACT  
 GCAAGAATTAAGGAAAAGGTGGAGAATGGAAGTGAACAGGGCCCTTACCCCCAGAG  
 CTGCAGCCTCTGTTGGAAGGGGAAGTAAAGGGGGCCCTGAGCCAACACCTCTGGTTC  
 AGACCTTTCTTAGACGCCAACAAAGGCTCCAGGAGATTGAACGCCGGCGAAATACTCCT  
 TTCCATGAGCGATTCTTCCGCTTCTTGTGGAAAGCTCTATGTCTTTCGCCGAGCTTC  
 CTGATGACTTGTATCTCACTTCAAACTGATATTAGGCCGCTCCTTCCCTGGAGCAGCT  
 GGCCAGGAGGTGACTTATGCAAACTTGAAGCCCTTGGAGCAGTTGGAGAAGTGAAT  
 CCCTCAAACACGGATTCTTACATTCTAATCCTCCTGAGTCAAATCCTGATCCTGTCCAC  
 TCAGAGTTCTGAAGGGGGCCAGATGTTGGAAGGGCAATTCGAGTCTAGAGGGCCCGCC  
 CTGATAGACGGTTTTTTCGCCCTTTCGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTT  
 GTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGAT  
 TTTGCCGATTTCCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAA  
 TTAATTCTGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG  
 CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAG  
 GCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTC  
 CCGCCCTAACTCCGCCATCCCGCCCTAACTCCGCCAGTTCCGCCATTCTCCGC  
 CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCTGCCTCTGAG  
 CTATTCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCCTAGGCTTTTGCAAAAGCTCCC  
 GGGAGCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCA  
 TGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATT  
 CGGCTGACTGGGCACAACAGACACTCGGCTGCTCTGATGCCGCCGTTTCCGGCTG  
 TCAGCGCAGGGCGCCCGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATG  
 AACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCTTGGC  
 CAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGT  
 GCCGGGGCAGGATCTCCTGTCTCATCTCACCTTGTCTCCTGCCGAGAAAGTATCCATCATG  
 GCTGATGCAATGCGGCGGCTGCATACGCTTGTCCGGCTACCTGCCATTGACCACC  
 AAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTGATCA  
 GGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGCCAGGCT  
 CAAGGCGCGCATGCCCGACGGCGAGGATCTCGTGTGACCCATGGCGATGCCTGCTT  
 GCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCTCGACTGTGGCCGGCTG  
 GGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGC  
 TTGGCGGCGAATGGGCTGACCGCTTCTCGTGCTTTACGGTATCGCCGCTCCCGATTC  
 GCAGCGCATCGCTTCTATCGCCTTCTTACGAGTTCTTCTGAGCGGGACTCTGGGGTT  
 CGAAATGACCGACCAAGCGACGCCAACCTGCCATCACGAGATTTGATTCCACCGCC  
 GCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCC  
 TCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCCACTTGTATTATGCAGCT  
 TATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCAC  
 TGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTATCATGTCTGTATACCGTC



GACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTA  
 TCCGCTCACAATCCACACAACATACGAGCCGGAAGCATAAAGTGAAAGCCTGGGGTG  
 CCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCG  
 GGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGGCGTT  
 TGCGTATTGGGCGCTCTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTTCGTT  
 GCTGCGGGGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGTTATCCACAGAATCA  
 GGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTA  
 AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA  
 AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG  
 TTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA  
 CCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGT  
 ATCTCAGTTCGGTGTAGGTGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT  
 TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC  
 ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT  
 AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACA  
 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC  
 TTGATCCGGCAAACAAACCACCGCTGGTAGCGTTTTTTTTGTTTGAAGCAGCAGATTA  
 CGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCT  
 CAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTC  
 ACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAAC  
 TTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATT  
 TCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCT  
 TACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGA  
 TTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTCTGCAACT  
 TTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGC  
 AGTTAATAGTTTGCACAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGT  
 CGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCC  
 CCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAA  
 GTTGGCCGCAAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCT  
 GCCATCCGTAAGATGCTTTTTCTGTGACTGGTGGTACTCAACCAAGTCATTCTGAGAATA  
 GTGTATGCGGCGACCGAGTTGCTCTTGCCCGCGTCAATACGGGATAATACCGCGCA  
 CATAGCAGAACTTTAAAGTGCTCATCATTGAAAACGTTCTTCGGGGCGAAAACCTCTCA  
 AGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATC  
 TTCAGCATCTTTACTTTACCCAGCGTTTCTGGGTGAGCAAACAGGAAGGCAAATG  
 CCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCTTTTTT  
 AATATTATTGAAGCATTTATCAGGGTATTGTCTCATGAGCGGATACATTTTGAATGTAT  
 TTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACG  
 TC //

SEQ ID NO: 13

Secuencia PP6

//gacggatcgggatctcccgatcccctatggtgcacictcagtacaatctgctctgatgccgatagtaagccagtatctgctc  
 cctgcttgtgttgaggctgctgagtagtgcgagcaaaatlaagctacaacaaggcaaggctgaccgacaattgcatga  
 agaatctgcttagggtagcggtttgctgctcgcgatgacgggacagatatacgcgttgacattgatttagctagtattaata  
 gtaatcaattacgggctcattagttcatagcccatataggagtccgcgttacataactacggtaaatggccgcctggctgaccg  
 cccaacgacccccgccattgacgtcaataatgacgtatgtcccatagtaacgccaatagggactttccattgacgtcaatgggt  
 ggagatttacggtaaaactgccactggcagtacatcaagtgtatcatatgccaagtagccccctattgacgtcaatgacggtaa  
 atggccgcctggcattatgccaglacatgacctatgggactttcctactggcagtlactctacgtattagctatgctattaccat  
 gggtatcggtttggcagtlactcaatgggctggatagcggttgactcacgggatttccaagctccacccattgacgtcaat  
 gggagttgtttggcaccaaaatcaacgggactttccaaatgtcgtacaactccgccccattgacgcaatgggctgtaggc

ggtacgggtgggaggtctatataagcagagctctctggctaactagagaaccactgcttactggcttatcgaaattaatacgactc  
 actatagggagaccaagctggctagcgtttaacttaagcttggtaccgagctcggatccactagtcagtggtggaattgcc  
 ctattcctcagtgaccaggagctgacacactatggcgacgctccgaggcttcagctgcctggctgctggccctggctgcct  
 ggtagccttgacacagccagcatgtgtcctgctcctcagcaagcagcgtcgtgctccagcgggtccggcgagccaacacc  
 ttctggaggaggtgctcaagggcaacctggagcagagtgctggaggagacgtgcagctacgaggaggccttcgaggctc  
 tggagtctccacggctacggatgtgttctgggccaagtacacagctgtgagacagcaggacgcctcgagataagctgctgc  
 atgtctggaaggaactgtgctgagggctgggtacgaactaccgagggcatgtgaacatcaccgggcaggcattgagtgcca  
 gctatggaggagctgctaccacataagcctgaaatcaactccactaccatcctggggccgacctacaggagaattctgccg  
 caaccccgacagcagcaccacgggaccctgggtctacactacagacccaccgtgaggaggcaggaatgcagcatccctgt  
 ctgtggccaggatcaagctactgtagcgtactccacgctccgaaggctccagtgtaactgtcacctccattggagcagtggt  
 ccctgatcgggggacagcagtaaccagggcgccctggcggtgaccacacatgggctcccctgctggcctgggcccagcgcacag  
 gccaaggccctgagcaagcaccaggactcaactcagctgtgcagctggggagaactctgcccgaacccagacgggggatg  
 aggagggcgtgtgtgctatgtggccgggaagcctggcgactttgggtactgcgacctcaactattgtgaggaggccgtggagg  
 aggagacaggagatgggtggtgaggactcagacagggccatcgaagggcgtaccgccaccagtgagtaaccagactttct  
 caatccgaggacctttggctcgggagaggcagactgtgggtcgcacctgtctgagaagaagtcgctggaggacaaaaccg  
 aaagagagctcctggaatcctacatcgacgggctggtggaggctcggatgcagagatcggcatgtcacctggcagggtga  
 tctttccggaagagtccccaggagctgctgtgtggggccagcctcatcagtgaccgctgggtcctaccgcccaccactgcctc  
 ctgtaccgcccctgggacaagaactcaccgagaatgaccttgggtgctgattggcaagcactcccgcaccaggtagcagcg  
 aaacattgaaaagatatccatgttgaaaagatctacatccacccaggtaacaactggcgggagaacctggaccgggacattg  
 ccctgatgaagctgaagaagcctgttcctcagtgactacattcacctgtgtctgcccgacagggagacggcagccagctg  
 ctccaggctggatacaagggcggggtgacagggctggggcaacctgaaggagacgtggacagccaacgttggaaggggca  
 gccagtgctcctgcaggtggtgaacctgccattgtggagcggccggctcgaaggactccaccggatccgcatcactgacaa  
 catgttctgtctggttacaagcctgatgaagggaaacgaggggatgctgtgaaggtagcagtgggggaccctttgcatgaag  
 agccctttaacaaccgctggtatcaaattggcatcgtctcatggggtaaggctgtgaccgggatgggaaataggtcttacac  
 acatgtgtccgctgaagaagtggtacagaaggtcattgatcagttggagagtagaagggcaattctgcagatataccagcac  
 agtggcgccgctcggctcctagagggcccgtttaaacccgctgatcagcctcactgtgcctctagtccagccatctgtgttg  
 cccctccccctgctccttgcacctggaagggtccactcccactgctccttctaataaaatgaggaaattgcatcgattgtctga  
 gtaggtgtcattctattctgggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggg  
 gatgcggtgggctctatggcttctgaggcggaaagaaccagctgggctctaggggtatcccacgcgccctgtagcggcgc  
 attagcggcggcggtgtgtggttacgcgcagcgtgaccgctacactgcccagcggcctagcggcctccttctgcttctccct  
 tcttctcggcacgttcggcgttccccgtcaagctctaaatcgggggctcccttagggttccgatttagtcttacggcacctcg  
 acccaaaaaactgattagggatggttccatcagatgcaatttctcattttattaggaaaggacagtgaggagtgccacctcc  
 aggttcaaggaaggcacgggggaggggcaaacacagatggctggcaactagaaggcacagtcgaggctgatcagcggg  
 tttaaacgggcccctctagactcgaattgccctccaacatctggccccctcagaactctgagtgagcaggatcaggattgactca  
 ggaggattagaatgtgaagaatccgtgttgaggattcagttctcaactgcctcaaagggtctcaagttgcataagtcacctcct  
 gggccagctgctccaggggaaggacggcctaataatcagatttcaagtgagatacaagtcacaggaagctcggcgaaagac

atagagcttgcacaagaagcggaagaatcgctcatggaaaggagtattcgccggcgtcaatctcctggagccttgttggc  
 gtctaagaaaggctcgaaccagagggtgtggctcagggcccccttactccccctccaacagaggctgcagctctgggggtaga  
 ggccctgttccattctccacctttccttaattctgagatagccaggctgtcctcagtgcaagctctgtagtgtagatag  
 acgtacatgtagcaagaagggtaggtgatgctgatacacctatggactaccagcaggcaactgcafttttccctctcgaa  
 gagtctggttctctgttctgccacaagctccacagtcactccccctgcagcagctggatgctagtggtgccaggcttcaactaca  
 aaattctccaagtcagtcagggaatctgcaatgaagaccacctcagtggtgtctaggctgctctgattcctgtaactggc  
 cctccagggagacagggtcaatcaagagtggtgacccaggatgtgctgaaagggtagcaagcggcctgcacgatgctc  
 acacgagggtaaaaaatcctctgctggaagcggctattgatggagacccaaatatcaaagtagatctggggctcagtgacattat  
 acttgggaagcaggcggctcaggcaagtggcatattgctcagcatgtctgcatgactcctccatcggcactctgtgaaataccc  
 cagggttaaggtagcccagtcgcccagtcgcccacacggtaggtgatctcagctgctggtgggagcgggagtgaccatcat  
 gtcccaggaatagccatacagcccatttgcagttgtatagccctgggtgagaaaaatgagaatagggcaggaatagctgctcc  
 agggagtagagcagggtgaaggcagctcccagctgatggcgcagccctggtctgcccactttgccccggctcctctataca  
 cacaggaaacactgggctgaggggctgcttggggcaacagttgtgcaacctcgggggagtaggacaccagcttccga  
 ggccactcaggggagcagaagagaggggctgctggccagcatgacgtaggagaacataccaatgctgaaaagctgggaattc  
 atgcatggaagtaggacacaaaagaacaggccaatggatctgagacatcaaaaagagcaggaaccagctgagaggctc  
 aagcagcagccccaccagtcacgaccagcaggctagtctcagctcctcagacaacagcagttgaagggactgaagagcca  
 gtccgggacaaaatattccatggaatagcctcaaccagctgcatccagcttttcaacccgcaatgaagtacacaatgaaga  
 tctggccacgggagcactgcatagtccaaaggggacgctgggctcctctctatgggcatcagcagaccgtccaagaccagt  
 agtgggttgcacatgaatgttagctgaaaggccaacaaccatacagataggagtggttccatgatgtctgtccaggagaa  
 acacataccagtagggcagcaggaataacacacagctatccggtagcacaggcccagcatcatgccagtgccccagaaa  
 catgatgggtgtagacaagatacatccagtcagtggtggcagtgggcgtagggcatccagcaaggggaagcggcacacatccagc  
 ccatcaaggtatttccgtccagagagctgagccccgctcctggggaatgcttagcaccatcaagaaccccaaaaagaaaacg  
 aaagacagctaagctgcagggtccgttggctgattcagcagggtcaccagcctccgcaactggacaaaatctgtccactcaaa  
 acccaagagttccctattcggctgctcctgctggccctgagatcagttcagcctgtcttctgactttatctgagctggcgaggctc  
 cgcgcggaaccggcagacaccgcccattgctctgcgaaagggcaattccaccacactggactagtggtccgagctcgttacc  
 aagcttaagttaaacgctagccagcttgggtcctctatagtgagctgtaatttcgataagccagtaagcagtggttctctagtt  
 agccagagagcttttgcaaaagcctaggcctccaaaaagcctcctcactacttctggaatagctcagaggccgagggcgccct  
 cggcctctgcataaaaaaaattagtcagccatggggcggagaatgggcggaactgggcggaggtagggcgggatgg  
 gcgagtagggcgggactatgggtgctgactaattgagatgcatgctttgcatactctgctgctggggagcctggggactttcc  
 acaccctaactgacacacattccacagccggatcgatgtggccatcgccctgatagacgggttttgccttgcagctggagctc  
 acgttcttaatagtgactctgttccaaactggaacaacactcaaccctatctcggtctattctttgattataagggatttgcgattt  
 cggcctattggttaaaaaatgagctgatttaacaaaaatgacggaattaattctgtggaatgtgtgctcagttagggtgtggaaggt  
 cccaggtccccagcaggcagaagtagcaagcatgcatctcaattagtcagcaaccaggtgtggaaggtccccaggtccc  
 ccagcaggcagaagtagcaagcatgcatctcaattagtcagcaaccatagtcggccccctaacccgccccatcccccccta  
 actccggcagctccgcccattctccggccatggctgactaattttttatgacagaggccaggccgctctgctctgagctta  
 ttccagaagtagtagggaggttttggaggcctaggcctttgcaaaaagctccccgggagctgtatcatttccgatctgatca

agagacaggatgaggatcgtttcgcatgattgaacaagatggattgcacgcagggttccggccgcttgggtggagaggctatc  
 ggctatgactgggcacaacagacaatcggtgctctgatgccgctgtccggctgtcagcgcagggcgcccgggtcttttgc  
 aagaccgacctgtccgggtccctgaatgaactgcaggacgagcagcggctatcgtggctggccacgacgggctcctg  
 cgcagctgtgtcgcagctgtcactgaagcgggaagggactggctgtattgggcgaagtccggggcaggatctcctgtcatc  
 caccttgcctcgcgagaaagtatccatcatggctgatgcaatgccggcgtgcatacgttgatccggctacctgccattcgac  
 caccaagcgaaacatcgcatcgagcgcagcactcggatggaagccggctctgtcgcagcagcagcagcagcagcagc  
 atcaggggctcgcgcccagccgaactgtccagcgtcaaggcgcgcagcggcagcagcagcagcagcagcagcagcagc  
 gatgctgttgcgaataatcaggtggaatggcggctttctggattcagcactgtggcggctgggtgtggcggaccgctac  
 aggacatagcgttggctaccctgatattgtgaagagctggcggcgaatgggctgaccgctcctcgtgctttacggatcgccg  
 ctcccgattcgcagcgcacgcctctatgcctcttgacgagttctctgagcgggactctggggtcgaaatgaccgaccaagc  
 gacgcccacactgcatcacgagatttcgatccaccgcccctctatgaaaggttgggctcggaaatcgtttccgggacgccc  
 gctggatgacctccagcggggatctcatgctggagttctcggccacccttctatgaaaggttgggctcggaaatcgtttccgggacgccc  
 agcaatagcatcacaatttcacaaataaagcatttttctactgcattctagtgtggttgcctcaaacatcaatgatctatcatgct  
 tgtataccgtcgacctctagctagagcttggcgtaacatggtcatagctgttctctgtgaaattgttatccgctcacaattccacac  
 aacatacgagccggaagcataaagtgtaaagcctgggggtcctaattgagtgagcctaactcacattaattgctgtcgtcactgc  
 ccgcttccagtcgggaaacctgctgcccagctgcatatgaatcggccaacgcgggggagaggcgggttgcgtatggggc  
 ctctccgctcctcgtcactgactcgtcgcctcggctcggctcggcgcagcggatcagctcactcaaaaggcggtaatac  
 gttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaag  
 gccgctgtcgtggcgttttccataggtcggccccctgacgagcatcaaaaaatcgacgctcaagtcagaggtggcgaaac  
 ccgacaggactataaagataaccaggcgttccccctggaagctccctcgtgctcctcgttccgacctgcccgttaccggatac  
 ctgtccgcttctccctcgggaagcgtggcgttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgtcgtccaagc  
 tgggctgtgtgacgaacccccgctcagcccagccgctgcgcctatccggtaactatcgtctgagccaaccggtaagacac  
 gacttatcgccactggcagcagccactggtaacaggatagcagagcggatgtagggcgtgctacagagttctgaagtgt  
 ggccaaactacggctacactagaagaacagatattggtatctgcgctcgtcgaagccagttacctcggaaaaagagttgtagct  
 ctgatccggcaaaacaaccaccgctgtagcgggtttttgttgaagcagcagattacgcgcagaaaaaaggatctcaaga  
 agatccttgatctttctacgggtctgacgctcagtggaacgaaaactcacgtaagggttttggcatgagattcaaaaagg  
 atcttaccttagatcctttaaataaaatgaagtttaataatcaatcaaaagtataatgagtaaacctggctgacagttaccaatgct  
 aatcagtgaggcacctatctcagcgaatcttattcgttcatccatagttgcctgactccccgctgtagataactcagatacggg  
 agggcttacctctggcccagtgctgcaatgataccgcgagaccacgctcaccggctccagattatcagcaataaacagc  
 cagccggaagggccgagcgaagaagtggtcctgcaacttaccgctccatccagctattaattgttccgggaagctagagt  
 aagtagttcgcagtaaatagtttgcgaacgtgtgctgacagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 agctccggttcccaacgatcaaggcaggtfacatgatccccatggttgcaaaaaagcgggtagctcctcgtcctccgatcgtt  
 gtcagaagtaagttggccgagtgatcactcatggtatggcagcactgcataattctctactgcatgccatccgtaagatgcttt  
 ctgtgactggtgagtactcaaccaagtattcigagaatagtgatgcggcagccgagttgctcttgcggcgcaatacgggata  
 ataccgcccacatagcagaactttaaagtgctcatattggaacagttctcggggcgaaaaactcaaggatcttaccgctg  
 ttgagatccagttcagatgaaccactcgtgcacccaactgatctcagcatctttactttaccagcgttctgggtgagcaaaaac

aggaaggcaaaatgccgcaaaaaggaataagggcgacacggaatgtgaatactcactcttcttttcaatatttgaa  
gcattatcagggtattgtctcatgagcggafacataattgaatgtattagaaaaataacaaataggggtccgcgacattccc  
cgaaaagtccacctgacgtc//

SEQ ID NO: 14

Secuencia F9NopaA

gacggatcgggagatctcccgatcccctatgggacactctcagtaacaatctgctctgatgccgcatagtaagccagtatctgclcc  
ctgctgtgtgtggaggctcgtgagtagtgccgagcaaaatfaagctacaacaaggcaaggctgaccgacaattgcatgaa  
gaatctgcttagggttaggcgttttgcgctgctcgcgatgtacggccagatatacgcgttgacattgatttagctagttattaatag  
taatcaattacggggtcattagttcatagcccataataggagtcccggttacataactacggtaaatggcccgcctggctgaccgc  
ccaacgacccccgccattgacgtcaataatgacgtatgttccatagtaacgccaatagggacttccattgacgtcaatgggtg  
gagtattacggtaaaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgcccctattgacgtcaatgacggtaaa  
tggcccgctggcattatgccagtaacatgaccttattggacttctacttggcagtaacatctacgtatttagctacgtattaccatg  
gtagtcgggtttggcagtaacatgaggcgtggatagcgggttgactcacgggattccaagtctccacccattgacgtcaatg  
ggagttgtttggcaccaaaatcaacgggacttccaaaatgtcgtacaactccgcccattgacgcaaatggcggttaggcgt  
gtacgggtggaggctatataagcagagctctctggtaactagagaaccactgcttactggcttatcgaaataatagcactcac  
tataggggagaccaagctggctagcgtttaaacttaagcttggtagcagctcgatccactagtcagtggtggaattgccctt  
attcctcagtgaccaggagctgacacacttagaagggcaattctgcagatcccatgcagcgcgtgaacatgatcatggcaga  
atcaccaggcctcatcaccatctgccccttaggalatctactcagtgctgaatgtacagttttctgatgaaaacgccaacaaaa  
ttctgaatcggccaaagaggataattcaggtaaatggaagagttgtcaagggaaccttgagagagaatgatggaagaaaa  
gltagtttgaagaagcagcagagaagttttgaaaactgaagaacaactgaattttggaagcagatgttgatggagatcagt  
gltagtcacatccatgtttaaattggcggcagtgcaaggatgacattaattcctatgaatgttggtgtccctttggattgaaggaaa  
aacgtgaattagatgtaacatgtaacattaagaatggcagatgcgagcagttttgtaaaaatagtgctgataacaaggtggttgc  
cctgtactgagggatctgacttcagaaaaaccagaagctctgtaaccagcagtgccattccatgtggaagagtttctgtttcac  
aaacttcaagctcaccctgctgagactgttttctgatgtggactatgaaattctactgaagctgaaaccattttggataacatca  
ctcaaagcaccacatcatttaatgacttcactcgggtgttggtggagaagatgccaaccaggfcaattccctggcaggtgtttg  
aatggtaaagttgatgcattctgtgaggctctatcgtaatgaaaaatggattgtaactgctgccactgtgtgaaactggtgtaa  
aattacagttgtcgcaggatgaacataatattgaggagacagaacatacagagcaaaagcgaatgtgattcgaattattctcac  
cacaactacaatgcagctattaataagtacaacatgacattgcccttctggaactggacgaacccttagtgctaaaacagctacgt  
tacacctattgcatgtgacaaggaatacagcaacatctcctcaaattggatctggctatgtaagtggctggggaagggctctcc  
acaaagggagatcagcttagttctcagtaccttagagttccactgttgaccgagccacatgtcttcgatctacaagttcaccatc  
tataacaacatgtctgtgctgctccatgaaggaggtagagattcatgtcaaggagatagtgggggaccccatgttactgaagtg  
gaaggaccagtttctaactggaaltattagctgggtgaagagtgcaatgaaaggcaaatatggaatatataccaaggtatc  
ccggtatgcaactggattaaggaaaaaacaagctcacttaatgaaagatggattccaaggatccagcacagtgccggccg  
ctcagactagagggcccgttaaaccgctgatcagcctcagctgtccttctagttgccagccatctgtgtttgccctccccctg  
gccttctgaccctggaagggtgccactcccactccttcttaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattc

tattctgggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtggg  
 ctctatggcttctgagcggaagaaccagctggggctctaggggatccccacgcgccctgtagcggcgattaagcggcg  
 ggggtggtggttacggcgagcgtgaccgctacacttgccagcgccctagcggccgctccttctccttctccttctcgccac  
 gttcgccggcttccccgtcaagctctaaatcgggggctccctttagggttccgatttagtcttacggcaccttcgacccccaaaa  
 acttgattagggctgtggaatgtgtgtcagttaggggtgtggaagtccccaggctccccagcaggcagaagatgcaaagcatgc  
 atctcaattagtcagcaaccagggtggaagtccccaggctccccagcaggcagaagatgcaaagcatgcatctcaattagt  
 cagcaacctatgcccggccctaaactccgcccctccccggcctaaactccgcccagttccgcccattctccggccctatggctgacta  
 attttttattatgacagggccgaggccgctcggcctctgagctattccagaagtagtgaggaggctttttggaggcctaggccttt  
 gcaaaaagctctctggctaactagagaaccactgctactggctatcgaaattaatcagactcactatagggagaccaagct  
 ggctagcgtttaacttaagcttggaaccgagctcggaaccactagtcagtggtggaattgcccttccgagagcaatggcgg  
 tctctgcccgggtccgcgggacctcggccagctcagataaagtacagaaagacaaggctgaactgatctcagggcccaggca  
 ggacagccgaataggaaactcttgggtttgagtggaacagattgtccagttggcggaggctggtgacctgtgaatcgacca  
 acggacctgcaagcttagctgtcttctgttttgggttctgatggtgctagacattccccaggagcgggggctcagctctctgg  
 accggaaataccttgatgggctggatgtgtgcccgtccccctgctggatgcccacgcccactgcccacttgactggatgtatctgtc  
 tacaccatcatgttctggggcactgggcatgatgctgggctgtgctaccggataagctgtgtgtattctctgctccatactggtat  
 gttttctctggacaagacatcatggaacaaccactcctatctgtaggggtggtggccttcagctaacattcatggatgcaacca  
 ctactggctgtggacggctgctgcaatgccataggaggaatgcccacgtgccccttggaaactatgcagtgctccgtggccagat  
 ctctattgtgacttcttggggtgtgaaaagctggatgcagactgggtgaaaggctattccatggaatattgtcccggcactggc  
 tctcagtccttcaaactgctgtgtctgaggagctgactagcctgctggtcgtgactgggggtgggctgctgctgacctctcagctg  
 gtttctgctctttttgatgtctcaagatccattggcctgttctgtctacttccactgcatgaattcccagctttcagcattggtatgtt  
 ctctacgtcatgctggccagcagccctctctctgctccccctgagtgccctcggaagctggtgtcctactgccccgaagggtgca  
 caactgtgcccctcaaggcagccccctcagcccagtgttcctgtgtgataagaggagccggggcaaaagtggccagaagcca  
 gggctgcccctcagctgggagctgccttcaacctgctctacctctggagcagctattcctgcccctattctcattttctaccccagg  
 ctataacaactggacaaatgggctgtatggctattcctgggacatgatggtgactcccgtcccaccagcagctgaagatcacct  
 accgtatggccgactggcgaactgggctaccttaacctggggtatttacacagagtcggcgatggaaggatcatgcagaca  
 tctgaagcaatagccactgctgagccgctgctcccaagtataatgtcactgagccccagatctactttgatatttgggtctcc  
 atcaatgaccgcttccagcagaggattttgacccctgtgtggacatcgtgcaggccgcttgggtacccttccagcgacatcctgg  
 gtgcaaccactctgatggacctgtctccctggaggccaagttacaggaaatcaagagcagcctagacaaccacactgaggt  
 ggtcttcattgacagattcctggactgcactggagaatgtgtgagtgaaagacctgggcaacactagcatccagctgctgcaggg  
 ggaagtgactgtggagctgtggcagaacagaagaaccagactctcgagaggagaaaaaagcagttgacctgctggtgagt  
 accataagggtatagacatcaccttagccctctgtctacatgtactgtcaacaactacagagctgcaactggagcaagacc  
 tggcatalctgcaagaattaaaggaaaagggtggagaatggaagtgaacagggcctctacccccagagctgcagcctctgtg  
 gaaggggaagtaaaagggggccctgagccaacacctctggttcagaccttcttagacgccaacaaaggctccaggagattga  
 acgcccggcaataactccttccatgagcgattctccgctctgtgtgcaaaagctctatgtcttccggcagcttctgatgactgt  
 atctcactcgaaatctgatattagccgcttccctggagcagctggcccaggaggtagctatgcaaaactgagacctttgag  
 gcagttggagaactgaatccctcaaacacggattctcaacttaacctcctgagtcataatcctgatcctgtccactcagagttctg

aagggggccagatgttgaagggaattcgagctagagggcccgcctgatagacggttttcgcccttgacgttggagtccac  
gttcttaatagtgactctgttccaaactggaacaacacccaacctatcggctctattctttgattataagggatttgcgatttgcg  
gcctattggttaaaaaatgagctgattaacaaaaatlaacgcgaattaattctgtggaatgtgtcagttagggttggaaagtcc  
ccaggctcccagcaggcagaagatgcaaagcatgcatctcaattagtcagcaaccaggtgtgaaagtccccaggctcccc  
agcaggcagaagatgcaaagcatgcatctcaattagtcagcaaccatagtcggccctaaactccgcccattcccggccctaac  
tccgcccagttccgcccattctccgcccattgctgactaattttttattatgacagggccgaggccgctctgctctgagctattc  
cagaagtagttagggaggctttttggaggcctaggctttgcaaaaagctcccgggagctgtataccatttggatctgatcaag  
agacaggatgaggatcgttccgatgattgaacaagatggattgcacgcagggttccggccgctgggtggagaggctattcgg  
ctatgactgggcacacagacaatcggtgctctgatgccgccgtgtccggctgacagcgaggggcgcccgggtcttttgc  
gaccgacctgtccgggtccctgaatgaactgcaggacgaggcagcgggctatcgtggctggccacgacggcggttcttgcg  
cagctgtctcagcgttgcactgaagcgggaaggactggctgctattgggcaagtccggggcaggatctcctgtcatctca  
cctgtcctcggcagaaaglatccatcatggctgatgcaatgcggcggctgcatacgtgatccggctacctgccattcgacca  
ccaagcgaacatcgcatcgagcagcagctactcggatggaagccggtctgtgatcaggatgatctggacgaagagcatc  
agggctcgcgcccagccgaactgtccgaggctcaaggcgcgatcccgcaggcaggatctcgtcgtgacctatggcga  
tgctgttccgcaatatcatggtgaaatggccgttttctggattcatcactgtggccggctgggtgtggcggaccgctatcag  
gacatagcgttggctaccctgatattgctgaagagctggcggcgaatgggctgaccgcttctcgtgctttacggatcgcgctc  
ccgattcgcagcgcacgctctatcgcctcttgacgagttctctgagcgggactctggggtcgaaatgaccgaccaagcgac  
gcccacctgccatcacgagattcagattccaccgcccctctatgaaaggtgggcttcggaatcgtttccgggacgcccgtg  
gatgatcctccagcggggatctcatgctggagtcttcccccactgtttattgcagctataatggttacaataaagc  
aatagcatcacaatttcacaataaagcatttttctcagctctagttgtggttttccaaactcatcaatgtatctatcatgtctgt  
ataccgtcagcctctagctagagcttggcgtaatcatggctatagctgttctctgtgaaattgtatccgctcacaattccacacaac  
atacagccggaagcataaagttaaagcctgggtgcctaatgagtgagcctaactcacattaattgcttgcgctcactgcccg  
ctttccagtcgggaaacctgtcgtccagctgcattaatgaatcgccaacgcgcccgggagaggcgggttgcgtattggcgtct  
tccgcttctcgtcactgactcgtcgcctcggctcgttcggctcggcgcgagcggatcagctcactcaaaggcggtaatacggfta  
tccacagaatcaggggataacgcaggaaagaacatgtagcaaaaaggccagcaaaaaggccaggaaccgtaaaaaggcc  
gctgtgctgggttttccataggctccgccccctgacgagcatcacaataatcagcgtcaagtcagaggtggcgaacccg  
acaggactataaagataccaggcgtttccccctggaagctcctcgtcgtctcctgttccgacctgcccgttaccggatacctg  
tccgcttctcccctcgggaagcgtggcgttctcatagctcacgctgtaggtatctcagttcgggtaggctgctccaagctg  
ggctgtgtgcacgaacccccgctcagcccagccgctgcgcttaccggtaactatcgtcttgcgttcaacccggtaagacag  
actatcccactggcagcagccactggtaacaggattagcagagcggatgtaggcgggtctacagagttctgaagtgtg  
gcctaactcggctacactagaagaacagatattggtatctcgtcgtctgtaagccagttacctcggaaaaagagttggtagctc  
ttgatccggcaacaaaccaccgctggtagcgggtttttgttgcagcagcagattacgcgcagaaaaaaaggatctcaagaa  
gatccttgaatctttctacgggtctgacgctcagtggaacgaaaactcacgtaagggatttgggtcatgagattcaaaaaggat  
cttccactagatccttttaataaaaatgaagtttaatacaatcaaatgatatagaglaaactggctgacagttaccaatgctta  
atcagtgaggcacctatctcagcgtctgtctattcgttcatccatagttgcctgactccccgctgtagataactcagatcggga  
gggcttaccatctggccccagtgctgcaatgataccgagacccacgctcaccggctccagattatcagcaataaaccagcc

agccggaagggccgagcgcagaagtggctcgaactttatccgcctccatccagcttattaattgtgcccgggaagctagagta  
 agtagtgcgccaagtaataagtttgcgcaacgttggtgccattgctacaggcatcggtggtcacgctcgttggatggctcattca  
 gctccggttcccaacgatcaaggcgagttacatgatccccatgtgtgcaaaaaagcggtagctcctcggctccgatcggtg  
 cagaagtaagtggccgcaagttatcactcatggtatggcagcactgcataattctctactgcatgccatccgtaagatgctttct  
 gtgactggtgagtactcaaccaagtcattctgagaatagtgatgcgggcaccgagttgctctgcccggcgtcaatacgggataat  
 accgcgccacatagcagaactttaaaagtctcatcattgaaaacgttctcggggcgaaaactctcaaggatcttaccgctgtt  
 gagatccagttcgaatgaaccactcgtgcacccaactgatcttcagcatctttactttaccagcgtttctgggtgagcaaaaaa  
 ggaaggcaaaatgccgcaaaaaaggaataagggcgacacggaaatgtgaatactcactcttcttttcaatattatgaag  
 ctttatcagggattgtctcatgagcggatcacatattgaaatgattagaaaaataaacaataggggtccgcgcacattcccc  
 gaaaagtgccacctgacgtc//

SEQ ID NO: 15

F9hglx

gacggatcgggagatctccgatcccctatggtgcaactcagtaaatctgctctgatgccgatagtaagccagatctgctcc  
 ctgctgtgtgtggaggctgctgagtagtgcgagcaaaatgaactacaacaaggcaaggctgaccgacaattgatgaa  
 gaatctgcttagggtaggcgttttgcgctcctcgatgtacgggacagataacgcgttgacattgattattgactagtattaatag  
 taatcaattacggggtcattagttcatagcccatatagagttccggttacataactacggtaaatggcccgcctggctgaccgc  
 ccaacgacccccgccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg  
 gagtattacggtaaacgccactggcagtaacatcaagtgatcatatgcaaglacgccccctattgacgtcaatgacggtaaa  
 tggcccgcctggcattatgcccagtaacatgacctatgggacttctacttggcagtaacatctacgtattagtcacgtattaccatg  
 gtgatcgggtttggcagtaacatgaggcgtgtagcgggttactcacgggattccaagctccacccattgacgtcaatg  
 ggagttgtttggcaccaaaatcaacgggactttccaaatgctgaacaactccgccccattgacgcaaatgggcggtaggcgt  
 gtacgggtggagggtctataagcagagctcctggtaactagagaaccactgcttactggcttatcgaaattaatacactcac  
 tatagggagacccaagctggctagcgtttaaacttaagcttggtaccgagctcgatccactagtccagtggtggaattgccctt  
 attcctcagtgaccaggagctgacacactagaagggaattctgcagataccatgcagcgcgtgaacatgatcatggcagaa  
 tcaccaggcctcatcaccatctgccttttaggatactactcagtgctgaatgtacagttttctgatcatgaaaacgccaacaaaatt  
 ctgaatcggccaaagaggataaattcaggtaaattggaagagttgttcaagggaaccttgagagagaatgatggaagaaaag  
 ttagtttgaagaagcagcagaagttttgaaaactgaagaacaactgaattttggaagcagatgttgatggagatcaggtg  
 gagtccaatccatgtttaaaggcggcagttgcaaggatgacattaattcctatgaatgttggtgtcccttggattgaaggaaaga  
 actgtgaattagatgtaacatgtaacattaagaatggcagatgagcagttttgtaaaaatagtgctgataacaagggtgtgctc  
 ctgtactgagggatctgactgcagaaaaccagaagcctgtgaaccagcagtgccatttccatgtggaagagtttctgtttaca  
 aacttcaagctcaccgctgctgagactgttttctgatgtggactatgtaaattctactgaagctgaaaccattttggataacatcact  
 caaagcaccatcatttaagacttactcgggtgtgtgggagaagatgccaaaccagggtcaattcccttggcaggtgttttga  
 atggtaaagttgatgattctgtggaggctctatcgtaataaaaaatggattgaactgctgccactgtgtgaaactgggtgtaaa  
 attacagttgctcaggtgaacataatattgaggagacagaacatacagagcaaaaagcgaatgtgattcgaatattcctcacc  
 acaactacaatgcagctattaataagtaacaacctgacattgcccttctggaactggacgaaccttagtgctaaacagctacgtt



acacctattgcattgctgacaaggaatacacgaacatcttctcaaattggatctggctatgtaagtggctggggaagggcttcc  
 acaaagggagatcagctttagtctcagcttagagttccactgttgaccgagccacatgtctcgatctcaaaagttcaccatc  
 tataacaacatgttctgtgctggcttccatgaaggagtagagattcatgtcaaggagatagtgggggaccccatgttactgaagtg  
 gaagggaccagttcttaactggaattattagctgggggaagagtgcaatgaaaggcaaatatggaatatataccaaggtatc  
 ccggtatgtcaactggattaaggaaaaaacaagctcactaatgaaagatggattccaaggatccagcacagtggcgccg  
 ctcggttctagagggcccgtttaaaccgctgatcagcctcagctgtgccttctagttgccagccatctgttggccctccccgt  
 gcctcctgaccctggaaggtgccactcccactgtccttctcaataaaatgaggaaattgcatcgcatgtctgagtaggtgtcattc  
 tattctgggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgtggggatgagggtggg  
 ctctatggcttctgaggggaaagaaccagctggggctctaggggtatccccacgcccctgtagcggcgcatgaagcgggc  
 ggggtgtgtggttacgcgacagctgaccgctacactgcccagcgccttagcggcccctcttctgcttctccctccttctcgccac  
 gttcgccggcttccccgtcaagctcaaatcgggggctcccttaggggtccgatttagtcttacggcacctcgacccccaaaaa  
 ctgattaggggtgatggtcacatcgatgcaatttctctatttaggaaaggacagtgaggagtgccacctccagggtaaggaa  
 ggcaagggggaggggcaacaacagatggctggcaactagaaggcacagctgaggctgatcagcgggttaaaccgggccc  
 tctagactcgaattgccctccaacatctggccccctcagaactctgagtgagcaggatcaggattgactcaggaggattagaat  
 tgaagaatccgtgttgagggtcagttctcaactgcctcaaaaggtctcaagttgcataagtcacctctggccagctgtc  
 caggaagggacggcctaataatcagattcgaagttagatacaagtcacaggaagctgcgcgaaagacatagagcttctgca  
 acaagaagcggagaagaatcgctcatggaaaggagtatttcggcggtcaatctctggagccttgttggcgtctaagaaaggt  
 ctgaaccagaggtgtggtcagggcccccttactccccctcaacagaggctgcagctctggggtagaggcctgttctactt  
 ccattctccaccttcttaattcttcagalatgccaggctgtctccagtgcaagctctgtagtggatagacgtacatgtagca  
 agaagggctaggtgatgtctatacaccttatgtactaccagcaggcaactgcattttctccctctcgaagagctggttctctg  
 ttctgccacaagctccacagtcactccccctgcagcagctggatgctagtggtgccaggctctcactcacaattctcaagtg  
 agtccagggaaatctgcaatgaagaccctcagtggtgtctaggtgctcttgatttctgtaactggccctccagggagac  
 aggtccatcaagagtggtgcaccaggatgtgcgtgaaagggtagcaagcggcctgcacgatgtccacacgagggtaaa  
 aaatcctctgctggaagcggctattgatggagacccaaatatcaaaagtagatctggggctcagtgacattatactgggaagcag  
 gcggtcaggaagtgcatattgctcagcatgtctgcatgatccttccatcgccgactctgtgaaataccccagggtaaggta  
 gccagttcgccagtgcgccatcacggtaggtgatctcacgtgctggtgggagcgggagtgaccatcatgtcccaggaata  
 gccatacagcccattgtccagttgtatagcccgggtgagaaaatgagaatagggcaggaatagctgctccaggaggtagag  
 cagggtaagggcagctccagctgatggcgagccctggcttctggccactttgccccggctcctctatacacaggaaca  
 ctgggctgaggggctgctttaggggcaacagttgtgcaaccttgggggagtaggacaccagctccgaggccactcagg  
 ggagcagaagagagggctgctggccagcatgacgttaggagaacataccaatgctgaaaagctgggaattcatgcagtgga  
 gtaggacacaaagaacaggccaatggatctgagacatcaaaaaagagcaggaaccagctgagaggtaagcagcagc  
 ccacccagtgacgaccagcaggctagtcagctcctcagacaacagcagttgaagggactgaagagccagtgccgggac  
 aaatattccatggaalagcctcaaccagctgcatccagcttttccaccccgaatgaagtacacaatgaagatctggccacg  
 gagcactgcatagttcaaaggggacgtgggcttctctatgggcttccagcagaccgtccacagaccagtagtggttgca  
 tccatgaatgttagctgaaaggccaacaaccatacagataggagtggttccatgatgtctgtccaggagaacacatacca  
 gtagggcagcaggaataacacacagcttaccggtagcacaggcccagcatcatgcccagtgccccagaaacatgatggtgt

agacaagatacatccagtcaagtggcagtgggcgtagggcatccagcaaggggaagcggcacacatccagccatcaaggt  
 attccgggtccagagagctgagccccgctcctggggaatgtctagcaccatcaagaacccaaaaagaaaacgaaagacagc  
 taagctgcaggggtccgttggtcgaltcagcaggggtaccagcctccgccaactggacaaatctgtccactcaaaaccaagagt  
 ttccctattcggctgctcctgggcccctgagatcagctcagcctgtcttctgtactttatctgagctgggcgagggtccgcgaggacc  
 cggcagacaccgcatgtctcgcggaaagggcaattccaccacactggactagtgatccgagctcgggtaccaagcttaagtt  
 aaacgctagccagctgggtctccctatagtgagtcgattaattcgcataagccagtaagcagtggggttctctagttagccagaga  
 gcttttgcaaaagcctaggcctcaaaaaagcctcctcactactctggaatagctcagagggcggaggcggcctcggcctctgc  
 ataaataaaaaaaattagtcagccatggggcggagaatggggcggaaactgggagggttaggggaggatgggagggttag  
 gggcgggactatggttctgactaattgagatgcatgctttgcatactctgctgctggggagcctggggactttccacaccctaac  
 tgacacacattccacagccggatcgatggtggccatgcctgatagacggttttcggccttgacgttgaggtccacgttcttaaf  
 agtggactctgttccaactggaacaacactcaacctatctcggctattctttgattataagggattttgccgatttcggcctattg  
 gtaaaaaatgagctgatttaacaaaaatftaacggaattaatctgtggaatgtgtcagttagggttggaagtccccaggct  
 ccccgagcaggcagaaglatgcaaagcatgcatctcaattagtcagcaaccaggtgtgaaagtccccaggctccccagcagg  
 cagaagatgcaaagcatgcatctcaattagtcagcaaccatagctcccggcctaacctcggccatcccgcccctaacctcggccc  
 agttccgcccattctccgcccattggctgactaattttttattatgcagaggccgaggccgctctgctctgagctattccagaagt  
 agtgaggaggctttttggaggcctaggctttgcaaaaagctcccgggagctgtatatacattttcggtatctgatcaagagacagg  
 atgaggatcgttctgcatgattgaacaagatggattgacgcaggttctccggcggctgggtggagaggctattcggctatgactg  
 ggcacaacagacaatcggctgctctgatgcccgtgttccggctgacgcagggggcggcgggttcttttgcagaccgac  
 ctgtccgggtgcccgaatgaactgaggacgaggcagcggcgtatcgtggctggccacgacggggcgttcttctgagcagctgtg  
 ctgcagctgtcactgaagcgggaagggactggctgctattgggccaagtgccggggcaggatctcctgcatctcacctgtctct  
 gccgagaaaglatccatcatggtcgalgcaatgcggcggctgcatacgttgatccggctacctgcccattcgaccaccaagcg  
 aaacatcgcatcgagcagcagcactcggatggaagccggcttctgcatcaggatgatctggacgaagagcatcaggggct  
 cgcgccagccgaactgtccagggctcaaggcgcgatcccgcagggcagggatctcgtcgtgacctatggcgtgctgctgct  
 gccgaatcatggtgaaaaatggccgctttctggattcatgactgtggccggctgggtgtggcggaccgctatcaggacatag  
 cgttggctaccctgatattgtgaagagcttggcggcgaatgggctgaccgcttctcgtgctttacggatcgcgctcccattc  
 gcagcgcacgcctctatgccttctgacgagttctctgagcgggactctggggtcgaatgaccgaccaagcgacgccc  
 cctgccatcacgagattcagttccaccgcccctctatgaaaggtgggctcggaaatcgtttccgggacgcccggctggatgatc  
 ctccagcgcggggatctcatgctggagtctctgcccaccccaactgtttattgcagcttataatggttacaataaagcaatagca  
 tcacaaattcacaataaagcatttttactgacttctagttgtgtttgtccaaactcatcaatgtatcttcatctgtataccgctc  
 gacctctagctagagcttggcgtaatcatggtcatagctgttctgtgtgaaattgtaaccgctcacaattccacacaacatacag  
 ccggaagcataaagtgtaaagcctggggtgcctaagtagtgagctaaactcaattaattgcgttgcgctcactgcccgtttccagt  
 cgggaaacctgtcgtgccagctgattaatgaatcggccaacgcgcggggagaggcgggttgcgtattgggctcttccgcttcc  
 tcgctcactgactcgtcgcctcggctcgttccgctcggcgcagcggatcagctcactcaaggcggtaatacggtaaccacaga  
 atcaggggataacgcaggaaagaacatgtgagcaaaagccagcaaaaggccaggaaccgtaaaaaggccgctgtctg  
 gcgtttttccataggctccgccccctgacgagcatcaaaaaatcgacgctcaagtcagagggtggcgaacccgacaggact  
 ataaagataccaggcgtttccccctggaagctcccctcgtgcgctctcctgttccgaccctcggccttaccggatacctgtccgcttcc

tccctcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgtcctccaagctgggctgtgtgc  
 acgaacccccctcagcccagccgctgcgcttatccggaactatcgtcttgagtccaacccggaagacacgacttatcgcca  
 ctggcagcagccactggtaacaggattagcagagcgaggatgtagggcggtgctacagagttctgaagtggtggcctaactac  
 ggctacactagaagaacagtattggatctgctgctctgctgaagccagttacctcggaaaaagagtggtgtagctcttgatccggc  
 aaacaaaccaccgctggtagcgggtttttgttgcaagcagcagattacgctcagaaaaaaggatctcaagaagatccttgat  
 cttttacggggtctgacgctcagtggaacgaaaactcacgtaagggattttggtcatgagattatcaaaaaggatctcacctag  
 atccttttaaataaaaatgaagtttaaatcaatctaaagtatatagtagtaaaactggctgacagttaccaatgcttaacagtgag  
 gcacctatctcagcgtatctgtctatttcgtcatccatagttgctgactccccgctcgtgtagataactacgatacgggagggttacc  
 atctggccccagtgctcaatgataccgagaccacgctcaccggctccagattatcagcaataaaccagccagccggaa  
 gggccgagcgcagaagtggtcctgcaactttatccgctccatccagcttattaattgttccgggaagctagagtaagtagttcgc  
 cagtaatagtttgcgcaacggtgtgcatgctacagggcatcgtggtgacagctcgtcgttggtaggtcattcagctccggttc  
 ccaacgatcaaggcgagttacatgatccccatgttgcaaaaaagcgggttagctcctcggctccgatcgtgtcagaagta  
 agttggccgagtggtatcactcatggtatggcagcactgcataattcttactgtcatgccatccgtaagatgcttttctgtgactggt  
 gactacaaaccaagtcattctgagaatagtgatgcgcgaccgagttgctcttgcggcgcaatacgggataataccgctc  
 cacatagcagaacttaaaagtctcatcattggaaaacggtcttggggcgaaaactcgaaggatctaccgctgttgagatcca  
 gttcagtaaacctcgtgcacccaactgatctcagcatctttactttaccagcgttctgggtgagcaaaaacaggaaggca  
 aaatgccgcaaaaaagggaataagggcgacacggaaatgtgaataactcactcttcttttcaatattatgaagcatttatcag  
 ggtattgtctcatgagcggatacatattgaatgtattgaaaaataaacaataggggtccgctcacatttccccgaaaagtgc  
 cacctgacgtc//

LISTADO DE SECUENCIAS

- <110> AstraZeneca AB
- <120> Proteína
- 5 <130> 101165-1P WO
- <140> PCT/SE2004/001453
- <141> 12-10-2004
- <150> 0324044.7
- <151> 14-10-2003
- 10 <160> 15
- <170> PatentIn versión 3.2
- <210> 1
- <211> 7422
- <212> DNA
- 15 <213> Artificial
- <220>
- <223> Vector
- <400> 1

# ES 2 357 603 T3

gacggatcgg gagatctccc gatecccctat ggtcgcactct cagtacaatc tgctctgatg	60
ccgcatagtt aagccagtat ctgctccctg cttgtgtggt ggaggtcgct gagtagtgcg	120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc	180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt	240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata	300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc	360
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc	420
attgacgtca atgggtggac tatttacggt aaactgcca cttggcagta catcaagtgt	480
atcatatgcc aagtagccc cctattgacg tcaatgacgg taaatggccc gcctggcatt	540
atgccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca	600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg	660
actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc	720

aaaatcaacg ggactttcca aaatgtcgt acaactccgc cccattgacg caaatgggcg 780  
 gtaggcgtgt acgggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840  
 ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gctggctagt 900  
 taagcttggg accgagctcg gatccactag tccagtgtgg tggattgcc cttattcctc 960  
 agtgaccag gagctgacac actatggcgc acgtccgagg cttgcagctg cctggctgcc 1020  
 tggccctggc tgccctgtgt agccttgtgc acagccagca tgtgttctg gctcctcagc 1080  
 aagcacggtc gctgctccag cgggtccggc gagccaacac cttcttggag gaggtgcgca 1140  
 agggcaacct ggagcgagag tgcgtggagg agacgtgcag ctacgaggag gccttcgagg 1200  
 ctctggagtc ctccacggct acggatgtgt tctgggcca gtacacaget tgtgagacag 1260  
 cgaggacgcc togagataag cttgctgcat gtctggaagg taactgtgct gagggtctgg 1320  
 gtacgaacta ccgagggcat gtgaacatca cccggtcagg cattgagtgc cagctatgga 1380  
 ggagtcgcta cccacataag cctgaaatca actccactac ccatcctggg gccgacctac 1440  
 aggagaatth ctgccgcaac cccgacagca gcaccacggg accctgggtgc taaactacag 1500  
 accccaccgt gaggaggcag gaatgcagca tccctgtctg tggccaggat caagtactg 1560  
 tagcgatgac tccacgctcc gaaggctcca gtgtgaatct gtcacctca ttggagcagt 1620  
 gtgtccctga tcgggggag cagtaccagg ggcgcctggc ggtgaccaca catgggctcc 1680  
 cctgcctggc ctgggccagc gcacaggcca aggccctgag caagcaccag gacttcaact 1740  
 cagctgtgca gctggtggag aacttctgcc gcaaccaga cgggatgag gagggcgtgt 1800  
 ggtgctatgt ggccgggaag cctggcgact ttgggtactg cgacctcaac tattgtgagg 1860  
 aggcctgga ggaggagaca ggagatgggc tggatgagga ctacagcagg gccatcgaag 1920  
 ggcgtaccgc caccagttag taccagactt tcttcaatcc gaggacctt ggctcgggag 1980  
 aggcagactg tgggctgca cctctgttcg agaagaagtc gctggaggac aaaaccgaaa 2040  
 gagagctcct ggaatcctac atcgacgggc gcattgtgga gggctcggat gcagagatcg 2100  
 gcatgtcacc ttggcaggtg atgcttttcc ggaagagtcc ccaggagctg ctgtgtggg 2160  
 ccagcctcat cagtgaccgc tgggtcctca ccgccccca ctgcctcctg taaccgccct 2220  
 gggacaagaa cttcaccgag aatgacctc tgggtgcgat tggcaagcac tcccgcacca 2280

ggtacgagcg aaacattgaa aagatatcca tgttgaaaa gatctacatc caccccaggt 2340  
 acaactggcg ggagaacctg gaccgggaca ttgccctgat gaagctgaag aagcctgttg 2400  
 ccttcagtga ctacattcac cctgtgtgtc tgcccagacag ggagacggca gccagcttgc 2460  
 tccaggctgg atacaagggg cgggtgacag gctggggcaa cctgaaggag acgtggacag 2520  
 ccaacgttg taaggggcag cccagtgtcc tgcaggtggt gaacctgccc attgtggagc 2580  
 ggccggtctg caaggactcc acccggatcc gcatcaactga caacatgttc tgtgctggtt 2640  
 acaagcctga tgaagggaaa cgaggggatg cctgtgaagg tgacagtggg ggacccttg 2700  
 tcatgaagag cccctttaac aaccgctggt atcaaatggg catcgtctca tggggtgaag 2760  
 gctgtgaccg ggatgggaaa tatggcttct acacacatgt gttccgcctg aagaagtgga 2820  
 tacagaaggt cattgatcag tttggagagt agaagggcaa ttctgcagat atccagcaca 2880  
 gtggcgcccg ctcgagtcta gagggccgc ggttcgaagg taagcctatc cctaaccctc 2940  
 tectcgttct cgattctacg cgtaccggtc atcatcacca tcaccattga gtttaaacc 3000  
 gctgatcagc ctcgactgtg ccttctagtt gccagccatc tgttgtttgc ccctccccg 3060  
 tgccttcctt gaccctggaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa 3120  
 ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg ggggtgggtg gggcaggaca 3180  
 gcaaggggga ggattgggaa gacaatagca ggcatgctgg ggatgcggtg ggctctatgg 3240  
 cttctgaggc ggaaagaacc agctggggct ctagggggta tccccacgcg cctgtagcg 3300  
 gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg 3360  
 ccctagcgcc cgtccttctc gctttcttcc cttcctttct cgccacgttc gccggcttcc 3420  
 cccgtcaagc tctaaatcgg ggcattccct tagggttccg atttagtgct ttacggcacc 3480  
 tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggcatcg ccctgataga 3540  
 cggtttttcg ccctttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa 3600  
 ctggaacaac actcaaccct atctcgttct attcttttga tttataaggg attttgggga 3660  
 tttcggccta ttggttaaaa aatgagctga ttaacaaaa atttaacgcg aattaattct 3720  
 gtggaatgtg tgtcagttag ggtgtggaaa gtccccagcg tccccaggca ggcagaagta 3780

tgcaaagcat gcatctcaat tagtcagcaa ccagggtgtgg aaagtcccca ggetccccag 3840  
 caggcagaag tatgcaaagc atgcatctca attagtcagc aaccatagtc ccgccccctaa 3900  
 ctccgcccac ccgcccccta actccgcccga gttccgcccga ttctccgccc catgggtgac 3960  
 taattttttt tatttatgca gaggccgagg ccgcctctgc ctctgagcta ttccagaagt 4020  
 agtgaggagg ctttttttga ggcctaggct tttgcaaaaa gctcccggga gcttgatat 4080  
 ccattttcgg atctgatcaa gagacaggat gaggatcgtt tcgcatgatt gaacaagatg 4140  
 gattgcacgc aggttctccg gccgcttggg tggagaggct attcggctat gactgggcac 4200  
 aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgcag gggcgcccgg 4260  
 ttctttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac gaggcagcgc 4320  
 ggctatcgtg gctggccacg acgggcgctt cttgcgcagc tgtgctcgac gttgtcactg 4380  
 aagcgggaag ggactggctg ctattggcg aagtgccggg gcaggatctc ctgtcatctc 4440  
 accttgetcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg ctgcatacgc 4500  
 ttgatccggc tacctgcccga ttcgaccacc aagcgaatac tcgcatcgag cgagcacgta 4560  
 ctccgatgga agccggctct gtcgatcagg atgatctgga cgaagagcat caggggctcg 4620  
 cgccagccga actgttcgcc aggctcaagg cgcgcatgcc cgacggcgag gatctcgtcg 4680  
 tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc ttttctggat 4740  
 tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg ttggctaccc 4800  
 gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg ctttacggta 4860  
 tcgccgctcc cgattcgcag cgcacgcct tctatgcct tcttgacgag ttcttctgag 4920  
 cgggactctg gggttcgcga aatgaccgac caagcgcgc ccaacctgcc atcacgagat 4980  
 ttcgattcca ccgccgctt ctatgaaagg ttgggcttcg gaatcgttt ccgggacgcc 5040  
 ggctggatga tctccagcg cggggatctc atgctggagt tcttcgcca cccaacttg 5100  
 tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt cacaaataaa 5160  
 gcattttttt cactgcattc tagttgtggt ttgtccaaac tcatcaatgt atcttatcat 5220  
 gtctgtatac cgtcgacctc tagctagagc ttggcgtaat catggtcata gctgtttcct 5280  
 gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt 5340

aaagcctggg gtgcctaata agtgagctaa ctacacattaa ttgcgttgcg ctcaactgccc 5400  
gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg 5460  
agaggcgggt tgcgtattgg gcgctcttcc gcttcctcgc tcaactgactc gctgcgctcg 5520  
gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca 5580  
gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 5640  
cgtaaaaagg ccgcgttgct ggcgttttcc cataggetcc gccccctga cgagcatcac 5700  
aaaaatcgac gctcaagtca gaggtggcga aaccgcagag gactataaag ataccagggc 5760  
tttccccctg gaagctccct cgtgcgctct cctgttccga cctgcgcgt taccggatac 5820  
ctgtccgctt ttctcccttc ggaagcgtg gcgctttctc aatgctcacg ctgtaggtat 5880  
ctcagttcgg ttaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 5940  
cccgaccgct gcgccttacc cggtaactat cgtcttgagt ccaaccgggt aagacacgac 6000  
ttatcggcac tggcagcagc cactggtaac aggattagca gagcgaggta ttagggcgg 6060  
gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttgg 6120  
atctgcgctc tgetgaagcc agttacctc ggaaaaagag ttggtagctc ttgatccggc 6180  
aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga 6240  
aaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 6300  
gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 6360  
cttttaaat aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggctc 6420  
gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 6480  
tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 6540  
ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 6600  
ataaaccagc cagccggaag ggccgagcgc agaagtggc ctgcaacttt atccgcctcc 6660  
atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagttg 6720  
cgcaacgttg ttgccattgc tacaggcatc gtgggtgctac gctcgtcgtt tggtaggct 6780  
tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa 6840



aaagcggtta gctccttcgg tcctccgata gttgtcagaa gtaagttggc cgcagtgta 6900  
 tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 6960  
 ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 7020  
 agttgctctt gcccggcgtc aatacgggat aataccgccc cacatagcag aactttaaaa 7080  
 gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctggtg 7140  
 agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 7200  
 accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggataaagg 7260  
 gcgacacgga aatgttgaat actcactatc ttcctttttc aatattattg aagcatttat 7320  
 cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaata 7380  
 ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tc 7422

<210> 2

<211> 10139

<212> DNA

5 <213> Artificial

<220>

<223> Vector

<400> 2

gacggatcgg gagatctccc gateccctat ggtgcaactc cagtacaatc tgctctgatg 60  
 ccgcatagtt aagccagtat ctgctccctg cttgtgtggt ggaggtcgct gagtagtgcc 120  
 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180  
 ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240  
 gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300  
 tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360  
 cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420  
 attgacgtca atgggtggag tatttacggt aaactgccc cttggcagta catcaagtgt 480  
 atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540  
 atgcccagta catgacctta tgggactttc ctacttgcca gtacatctac gtattagtca 600  
 tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660

actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720  
 aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780  
 gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840  
 ctgcttactg gcttatcgaa attaatacga ctcaactatag ggagacccaa gctggctagc 900  
 gtttaaactt aagcttggtg ccgagctcgg atccactagt ccagtgtggt ggaattgccc 960  
 ttattcctca gtgaccagg agctgacaca ctatggcgca cgtccgaggc ttgcagctgc 1020  
 ctggctgcct ggccctggct gccctgtgta gccttgtgca cagccagcat gtgttctctg 1080  
 ctccctcagca agcacggctc ctgctccagc gggtcggcg agccaacacc ttcttgagg 1140  
 aggtgcgcaa gggcaacctg gagcgagagt gcgtggagga gacgtgcagc tacgaggagg 1200  
 ccttcgaggc tctggagtcc tccacggcta cggatgtgtt ctgggccaag tacacagctt 1260  
 gtgagacagc gaggacgcct cgagataagc ttgctgcatg tctggaaggt aactgtgctg 1320  
 agggctctggg tacgaactac cgagggcatg tgaacatcac ccggtcaggc attgagtgcc 1380  
 agctatggag gagtcgctac ccacataagc ctgaaatcaa ctccactacc catcctgggg 1440  
 ccgacctaca ggagaatttc tgccgcaacc ccgacagcag caccacggga ccctgggtgct 1500  
 aactacaga ccccaccgtg aggaggcagg aatgcagcat ccctgtctgt ggccaggatc 1560  
 aagtcactgt agcgatgact ccacgctccg aaggctccag tgtgaatctg tcacctccat 1620  
 tggagcagtg tgteccctgat cgggggcagc agtaccaggg gcgcctggcg gtgaccacac 1680  
 atgggctccc ctgcctggcc tgggcccagc cacaggccaa ggccctgagc aagcaccagg 1740  
 acttcaactc agctgtgcag ctggtggaga acttctgccg caaccacagc ggggatgagg 1800  
 agggcgtgtg gtgctatgtg gccgggaagc ctggcgactt tgggtactgc gacctcaact 1860  
 attgtgagga ggccgtggag gaggagacag gagatgggct ggatgaggac tcagacaggg 1920  
 ccatcgaagg gcgtaccgcc accagtgagt accagacttt ctcaatccg aggaccttg 1980  
 gctcgggaga ggcagactgt gggctgcgac ctctgttcga gaagaagtcg ctggaggaca 2040  
 aaaccgaaag agagctctg gaatcctaca tcgacgggcg cattgtggag ggctcggatg 2100  
 cagagatcgg catgtcacct tggcaggtga tgcttttccg gaagagtccc caggagctgc 2160

tgtgtggggc cagcctcatc agtgaccgct gggtoctcac cgccgccac tgccctcctgt 2220  
acccgccctg ggacaagaac ttcaccgaga atgaccttct ggtgcgcatt ggcaagcact 2280  
cccgaccag gtacgagcga aacattgaaa agatatccat gttggaaaag atctacatcc 2340  
accccaggta caactggcgg gagaacctgg accgggacat tgccctgatg aagctgaaga 2400  
agcctgttgc cttcagtac tacattcacc ctgtgtgtct gcccgacagg gagacggcag 2460  
ccagcttget ccaggetgga tacaaggggc gggtgacagg ctggggcaac ctgaaggaga 2520  
cgtggacagc caacgttggg aaggggcagc ccagtgtcct gcaggtgggtg aacctgccc 2580  
ttgtggagcg gccggtctgc aaggactcca cccggatccg catcactgac aacatgttct 2640  
gtgctggtta caagcctgat gaagggaaac gaggggatgc ctgtgaaggt gacagtgggg 2700  
gacctttgt catgaagagc cccttaaca accgctggta tcaaatgggc atcgtctcat 2760  
gggggtaagg ctgtgaccgg gatgggaaat atggcttcta cacacatgtg ttccgctga 2820  
agaagtggat acagaaggtc attgatcagt ttggagagta gaagggcaat tctgcagata 2880  
tccagcacag tggcggccgc tcgagtctag agggcccggt taaaccgct gatcagcctc 2940  
gactgtgcct tctagttgcc agccatctgt tgttgcccc tccccgtgc cttccttgac 3000  
cctggaagggt gccactccc ctgtcctttc ctaataaaat gaggaaattg catcgattg 3060  
tctgagtagg tgtcattcta ttctgggggg tggggggggg caggacagca agggggagga 3120  
ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgagggcga 3180  
aagaaccagc tggggctcta gggggtatcc ccacgcgcc tgtagcggcg cattaagcgc 3240  
ggcgggtgtg gtggttacgc gcagcgtgac cgtacactt gccagcgccc tagcggccgc 3300  
tcctttcgtc ttcttcctt cctttctcgc cacgttcgcc ggctttcccc gtcaagetct 3360  
aatcggggg ctcccttag ggttcgatt tagtgettta cggcaccttc gaccccaaaa 3420  
aacttgatta gggctgtgga atgtgtgtca gttaggggtg ggaaagtc caggctcccc 3480  
agcaggcaga agtatgcaaa gcatgcatct caattagtca gcaaccaggt gtggaaagtc 3540  
cccaggctcc ccagcaggca gaagtatgca aagcatgcat ctcaattagt cagcaacct 3600  
agtccccccc ctaactccgc ccatcccccc cetaactccg cccagttccg cccattctcc 3660  
gccccatggc tgactaattt tttttattha tgcagaggcc gagggccgct cggcctctga 3720

gctattccag aagtagtgag gaggcttttt tggaggccta ggcttttgca aaaagctctc 3780  
tggctaacta gagaaccac tgcttactgg cttatcgaaa ttaatacgac tcaactatagg 3840  
gagaccaag ctggctagcg tttaaactta agcttggtag cgagctcgga tccactagtc 3900  
cagtgtggtg gaattgcctt tccgcagag caatggcggg gtctgccggg tccgcgcgga 3960  
cctcgcaccag ctcagataaa gtacagaaa gacaaggctga actgatctca gggcccaggc 4020  
aggacagccg aatagggaaa ctcttggggtt ttgagtggac agatttgctc agttggcgga 4080  
ggctggtagc cctgctgaat cgaccaacgg accctgcaag cttagctgtc tttcgttttc 4140  
tttttggggtt cttgatggtg ctagacattc cccaggagcg ggggctcagc tctctggacc 4200  
ggaaatacct tgatgggctg gatgtgtgcc gcttcccctt gctggatgcc ctacgcccac 4260  
tgccacttga ctggatgat cttgtctaca ccatcatggt tctgggggca ctgggcatga 4320  
tgctgggect gtgctaccgg ataagctgtg tgttattcct gctgccatac tggatgtgt 4380  
ttctcctgga caagacatca tggaacaacc actcctatct gtatgggttg ttggcctttc 4440  
agctaacatt catggatgca aaccactact ggtctgtgga cggctctgctg aatgcccata 4500  
ggaggaatgc ccacgtgccc ctttggaaact atgcagtgct ccgtggccag atcttcattg 4560  
tgtacttcat tgccgggtgtg aaaaagctgg atgcagactg ggttgaaggc tattccatgg 4620  
aatatttgtc ccggcactgg ctcttcagtc ccttcaaact gctgttgtct gaggagctga 4680  
ctagcctgct ggtcgtgca c tggggtgggc tgctgcttga cctctcagct ggtttcctgc 4740  
tcttttttga tgtctcaaga tccattggcc tgttcttgt gtectacttc cactgcatga 4800  
attcccagct tttcagcatt ggtatgttct cctacgtcat gctggccagc agccctctct 4860  
tctgctcccc tgagtggcct cggaagctgg tgtcctactg cccccgaagg ttgcaacaac 4920  
tggtgcccct caaggcagcc cctcagcca gtgtttcctg tgtgtataag aggagccggg 4980  
gcaaaaagtgg ccagaagcca gggctgcgcc atcagctggg agctgccttc accctgctct 5040  
acctcctgga geagctattc ctgccctatt ctcattttct caccagggc tataacaact 5100  
ggacaaaatgg gctgtatggc tattcctggg acatgatggt gcaactcccgc tcccaccagc 5160  
acgtgaagat cacctaccgt gatggcgcga ctggcgaact gggetacctt aaccctgggg 5220

tatttacaca gagtcggcga tggaaggatc atgcagacat gctgaagcaa tatgccactt 5280  
gcctgagccg cctgcttecc aagtataatg tcaetgagcc ccagatctac tttgatattt 5340  
gggtctccat caatgaccgc ttccagcaga ggatttttga ccctcgtgtg gacatcgtgc 5400  
aggccgcttg gtcacccttt cagcgcacat cctgggtgca accactcttg atggacctgt 5460  
ctccctggag ggccaagtta caggaaatca agagcagcct agacaaccac actgaggtgg 5520  
tcttcattgc agattteccct ggactgcact tggagaattt tgtgagtga gacctgggca 5580  
aactagcat ccagctgctg cagggggaag tgactgtgga gcttgtggca gaacagaaga 5640  
accagactct tcgagaggga gaaaaaatgc agttgcctgc tggtagtac cataaggtgt 5700  
atagacatc acctagccct tcttgctaca tgtacgtcta tgtcaacact acagagcttg 5760  
cactggagca agacctggca tatctgcaag aattaaagga aaaggtggag aatggaagtg 5820  
aaacagggcc totacecca gagctgcagc ctctgttggga aggggaagta aaagggggcc 5880  
ctgagccaac acctctggtt cagaccttcc ttagacgcca acaaaggctc caggagattg 5940  
aacgccggcg aaatactcct ttccatgagc gattcttccg cttcttgttg cgaaagctct 6000  
atgtctttcg ccgcagcttc ctgatgactt gtatctcact tcgaaatctg atattaggcc 6060  
gtccttcctt ggagcagctg gcccaggagg tgacttatgc aaacttgaga ccctttgagg 6120  
cagttggaga actgaatecc tcaaacacgg attcttcaca ttctaactct cctgagtcaa 6180  
atcctgatec tgtccactca gagttctgaa gggggccaga tgttgaagg gcaattcgag 6240  
tctagagggc ccgccctgat agacggtttt tcgccctttg acgttggagt ccacgttctt 6300  
taatagtgga ctcttgttcc aaactggaac aacactcaac cctatctcgg tctattcttt 6360  
tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc tgatttaaca 6420  
aaaatttaac gcgaattaat tctgtggaat gtgtgtcagt tagggtgtgg aaagtcccca 6480  
ggctccccag caggcagaag tatgcaaagc atgcactca attagtcagc aaccaggtgt 6540  
ggaaagtccc caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca 6600  
gcaaccatag tcccgccctt aactccgcc atccccccc taactccgcc cagttccgcc 6660  
cattctccgc cccatggctg actaattttt tttatttatg cagagggcca ggccgcctct 6720  
gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg cttttgcaaa 6780

aagctcccgg gagcttgtat atccattttc ggatctgate aagagacagg atgaggatcg 6840  
tttcgcatga ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg 6900  
ctattcggct atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg 6960  
ctgtcagcgc agggggcggc ggttcttttt gtcaagaccg acctgtccgg tgccctgaat 7020  
gaactgcagg acgaggcagc gggctatcg tggctggcca cgacgggctg tccttgcgca 7080  
gctgtgctcg acgttgtcac tgaagcggga agggactggc tgctattggg cgaagtgccg 7140  
gggcaggatc tcctgtcadc tcaccttgcct cctgccgaga aagtatccat catggctgat 7200  
gcaatgcggc ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaaa 7260  
catcgcatcg agcagcagc tactcggatg gaagccggtc ttgtcgatca ggatgatctg 7320  
gacgaagagc atcaggggct cgcgccagcc gaactgttcg ccaggctcaa ggcgcgcatg 7380  
cccgaaggcg aggatctcgt cgtgacccat ggcgatgcct gcttgccgaa tatcatggtg 7440  
gaaaatggcc gcttttctgg attcatcgac tgtggccggc tgggtgtggc ggaccgctat 7500  
caggacatag cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac 7560  
cgcttcctcg tgctttacgg tatcgccgct cccgatcgc agcgcacgc cttctatcgc 7620  
cttcttgacg agttctctg agcgggactc tggggttcga aatgaccgac caagcgacgc 7680  
ccaacctgcc atcaagagat ttogattcca ccgcgcctt ctatgaaagg ttgggcttcg 7740  
gaatcgtttt ccgggacgcc ggttggatga tcctccagcg cggggatctc atgctggagt 7800  
tcttcgcca cccaacttg ttattgcag cttataatgg ttacaataa agcaatagca 7860  
tcacaaattt cacaaataa gcattttttt cactgcattc tagttgtggt ttgtccaaac 7920  
tcatcaatgt atcttatcat gtctgtatac cgtcgacctc tagctagagc ttggcgtaat 7980  
catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac 8040  
gagccggaag cataaagtgt aaagcctggg gtgcctaag agtgagctaa ctacattaa 8100  
ttgcgttgcg ctactgcc gctttccagt cgggaaacct gtcgtgccag ctgcattaat 8160  
gaatcgcca acgcgcgggg agaggcgggt tgcgtattgg gcgctcttcc gcttcctcgc 8220  
tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 8280

cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 8340  
 gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggetcc 8400  
 gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 8460  
 gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 8520  
 ccctgccgct taccggatac ctgtccgctt ttctccctc gggaagcgtg gcgctttctc 8580  
 atagctcacg ctgtaggtat ctcaagtccg tgtaggtcgt tcgctccaag ctgggctgtg 8640  
 tgcacgaacc ccccgttcag cccgaccgct gcgccttata cggtaactat cgtcttgagt 8700  
 ccaaccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 8760  
 gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 8820  
 ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttacctc ggaaaaagag 8880  
 ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggttttttt gtttgcaagc 8940  
 agcagattac gcgcagaaaa aaaggatctc aagaagatcc ttgatcttt tctacggggt 9000  
 ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 9060  
 ggatcttcac ctagatcctt ttaaattaa aatgaagttt taaatcaatc taaagtatat 9120  
 atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 9180  
 tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac 9240  
 gggagggcct accatctggc cccagtgtg caatgatacc gcgagacca cgctcacccg 9300  
 ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggctcctg 9360  
 caactttatc cgcctccatc cagtctatta attggtgccc ggaagctaga gtaagtagtt 9420  
 cgccagttaa tagtttgccg aacgttgttg ccattgctac aggcatcgtg gtgtcacgct 9480  
 cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 9540  
 ccccatggt gtgcaaaaaa gcggttagct ccttcggctc tccgatcgtt gtcagaagta 9600  
 agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca 9660  
 tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat 9720  
 agtgatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcccac 9780  
 atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 9840

ggatcctacc gctggtgaga tccagttcga tgtaaccac tcgtgcaccc aactgatcct 9900  
 cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg 9960  
 caaaaaaggg aataaggcg acacggaaat gttgaatact catactcttc ctttttcaat 10020  
 attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 10080  
 agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgccca cctgacgctc 10139

<210> 3  
 <211> 26  
 <212> DNA  
 5 <213> Homo sapiens  
 <400> 3  
 attcctcagt gaccaggag ctgaca 26  
 <210> 4  
 <211> 42  
 10 <212> DNA  
 <213> Homo sapiens  
 <400> 4  
 ctactctcca aactgatcaa tgacctctg tatccacttc tt 42  
 <210> 5  
 15 <211> 23  
 <212> DNA  
 <213> Homo sapiens  
 <400> 5  
 tccgcagagc aatggcggtg tct 23  
 20 <210> 6  
 <211> 24  
 <212> DNA  
 <213> Homo sapiens  
 <400> 6  
 25 ccaacatctg gcccttcag aact 24  
 <210> 7  
 <211> 24  
 <212> DNA  
 <213> Homo sapiens  
 30 <400> 7  
 acacctctgg ttcagacctt tctt 24



<210> 8  
 <211> 24  
 <212> DNA  
 <213> Homo sapiens  
 5 <400> 8  
 aatcgctcat ggaaaggagt attt 24  
 <210> 9  
 <211> 27  
 <212> DNA  
 10 <213> Homo sapiens  
 <400> 9  
 caacaaaggc tccaggagat tgaacgc 27  
 <210> 10  
 <211> 22  
 15 <212> DNA  
 <213> Homo sapiens  
 <400> 10  
 tggaggacaa aaccgaaaga ga 22  
 <210> 11  
 20 <211> 18  
 <212> DNA  
 <213> Homo sapiens  
 <400> 11  
 catccgagcc ctccacaa 18  
 25 <210> 12  
 <211> 25  
 <212> DNA  
 <213> Homo sapiens  
 <400> 12  
 30 ctcttgaat cctacatcga cgggc 25  
 <210> 13  
 <211> 10238  
 <212> DNA  
 <213> Artificial  
 35 <220>  
 <223> Vector  
 <400> 13

ES 2 357 603 T3

gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60  
ccgcatagtt aagccagtat ctgctccctg cttgtgtggt ggaggtcgct gagtagtgcg 120  
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180  
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240  
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300  
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360  
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420  
attgacgtca atgggtggag tatttacggg aaactgcca cttggcagta catcaagtgt 480  
atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540  
atgccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600  
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660  
actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc 720  
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780  
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840  
ctgcttactg gcttatcgaa attaatacga ctcaactatag ggagaccaa gctggctagc 900  
gtttaaactt aagcttggtg ccgagctcgg atccactagt ccagtgtggt ggaattgcc 960  
ttattcctca gtgaccagg agctgacaca ctatggcgca cgtccgaggc ttgcagctgc 1020  
ctggctgect ggccctggct gccctgtgta gccttgtgca cagccagcat gtgttcctgg 1080  
ctcctcagca agcacggtcg ctgctccagc gggtcggcg agccaacacc ttcttgagg 1140  
aggtgcgcaa gggcaacctg gagcgagagt gcgtggagga gacgtgcagc tacgaggagg 1200  
ccttcgaggc tctggagtcc tccacggcta cggatgtggt ctgggccaag tacacagctt 1260

ES 2 357 603 T3

gtgagacagc gaggacgcct cgagataagc ttgctgcatg tctggaaggt aactgtgctg 1320

agggtctggg tacgaaactac cgagggcatg tgaacatcac ccggtcaggc attgagtgcc 1380

agctatggag gagtcgctac ccacataagc ctgaaatcaa ctccactacc catcctgggg 1440

ccgacctaca ggagaatttc tgccgcaacc ccgacagcag caccacggga ccctgggtgct 1500

acactacaga ccccaccgtg aggaggcagg aatgcagcat ccctgtctgt ggccaggatc 1560

aagtcactgt agcgatgact ccacgctccg aaggctccag tgtgaatctg tcacctccat 1620

tggagcagtg tgtccctgat cgggggcagc agtaccaggg gcgcctggcg gtgaccacac 1680

atgggctccc ctgcctggcc tgggccagcg cacaggccaa ggccctgagc aagcaccagg 1740

acttcaactc agctgtgcag ctggtggaga acttctgccg caaccagac ggggatgagg 1800

agggcgtgtg gtgctatgtg gccgggaagc ctggcgactt tgggtactgc gacctcaact 1860

attgtgagga ggccgtggag gaggagacag gagatgggct ggatgaggac tcagacaggg 1920

ccatcgaagg gcgtaccgcc accagtgagt accagacttt cttcaatccg aggacctttg 1980

gctcgggaga ggcagactgt gggctgcgac ctctgttcca gaagaagtcg ctggaggaca 2040

aaaccgaaag agagctcctg gaatcctaca tcgacgggcg cattgtggag ggctcggatg 2100

cagagatcgg catgtcacct tggcaggtga tgcttttccg gaagagtccc caggagctgc 2160

tgtgtggggc cagcctcatc agtgaccgct gggctctcac cgccgccac tgctcctgt 2220

accgcctctg ggacaagaac ttcaccgaga atgaccttct ggtgcgcatt ggcaagcact 2280

ccgcaccag gtacgagcga aacattgaaa agatatccat gttggaaaag atctacatcc 2340

accccaggta caactggcgg gagaacctgg accgggacat tgccctgatg aagctgaaga 2400

agcctgttgc cttcagtgac tacattcacc ctgtgtgtct gcccgacagg gagacggcag 2460

ccagcttgct ccaggctgga tacaaggggc gggtgacagg ctggggcaac ctgaaggaga 2520

cgtggacagc caacgttggt aaggggcagc ccagtgtcct gcagggtggtg aacctgccca 2580

ttgtggagcg gccggtctgc aaggactcca cccggatccg catcactgac aacatgttct 2640

gtgctggtta caagcctgat gaagggaaaac gaggggatgc ctgtgaaggt gacagtgggg 2700

gaccctttgt catgaagagc ccctttaaca accgctggta tcaaatgggc atcgtctcat 2760

ggggtgaagg ctgtgaccgg gatgggaaat atggcttcta cacacatgtg ttccgcctga 2820

agaagtggat acagaaggtc attgatcagt ttggagagta gaagggcaat tctgcagata 2880  
 tccagcacag tggcggccgc tcggttccta gagggcccgt ttaaaccgc tgatcagcct 2940  
 cgactgtgcc ttctagttgc cagccatctg ttgtttgccc ctccccgtg ccttccttga 3000  
 ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt gcatcgcatt 3060  
 gtctgagtag gtgtcattct attctggggg gtgggggtggg gcaggacagc aagggggagg 3120  
 attgggaaga caatagcagg catgctgggg atgcggtggg ctctatggct tctgaggcgg 3180  
 aaagaaccag ctggggctct agggggatc cccacgcgcc ctgtagcggc gcattaagcg 3240  
 cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcctg 3300  
 ctcttttcgc tttcttcctt tctttctcg ccacgttcgc cggctttccc cgtcaagctc 3360  
 taaatcgggg gctcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa 3420  
 aacttgatta gggatgatgt tcacatcgat gcaatttctt cattttatta ggaaaggaca 3480  
 gtgggagtgg caccttcacg ggtcaaggaa ggcacggggg aggggcaaac aacagatggc 3540  
 tggcaactag aaggcacagt cgaggctgat cagcgggttt aaacgggccc tctagactcg 3600  
 aattgccctt ccaacatctg gccccctca gaactctgag tggacaggat caggatttga 3660  
 ctcaggagga ttagaatgtg aagaatccgt gtttgagga ttcagttctc caactgcctc 3720  
 aaagggcttc aagtttgc atagtcacctc ctgggcccagc tgctccaggg aaggacggcc 3780  
 taatatcaga tttcgaagtg agatacaagt catcaggaag ctgcggcgaa agacatagag 3840  
 ctttcgcaac aagaagcggg agaatcgctc atggaaagga gtatttcgcc ggcgttcaat 3900  
 ctctgggagc ctttggtggc gtctaagaaa ggtctgaacc agaggtgttg gctcagggcc 3960  
 cccttttact tccccttcca acagaggctg cagctctggg ggtagaggcc ctgtttcaet 4020  
 tccattctcc accttttctt ttaattcttg cagatatgcc aggtcttgct ccagtgcaag 4080  
 ctctgtagtg ttgacataga cgtacatgta gcaagaaggg ctaggtgatg tcgtatacac 4140  
 cttatggtac tcaccagcag gcaactgcat tttttctccc tctcgaagag tctggttctt 4200  
 ctgttctgcc acaagctcca cagtcacttc ccctgcagc agctggatgc tagtgttgcc 4260  
 caggcttcca ctacaaaat tctccaagtg cagtcagggg aaatctgcaa tgaagaccac 4320

ctcagtgtgg ttgtctaggc tgctcttgat ttctgtaac ttggccctcc agggagacag 4380  
gtccatcaag agtggttgca cccaggatgt gcgctgaaag ggtgaccaag cggcctgcac 4440  
gatgtccaca cgaggggtcaa aaatcctctg ctggaagcgg tcatgatgg agacccaaat 4500  
atcaaagtag atctggggct cagtgcatt atacttggga agcaggcggc tcaggcaagt 4560  
ggcatattgc ttcagcatgt ctgcatgac cttccatcgc cgactctgtg taaatacccc 4620  
agggttaagg tagcccagtt cgccagtgcg gccatcacgg taggtgatct tcacgtgctg 4680  
gtgggagcgg gagtgcacca tcatgtccca ggaatagcca tacagcccat ttgtccagtt 4740  
gttatagccc tgggtgagaa aatgagaata gggcaggaat agctgctcca ggaggtagag 4800  
cagggtgaag gcagctccca gctgatggcg cagccctggc ttctggccac ttttgccccg 4860  
gctcctotta tacacacagg aaacactggg ctgaggggct gccttgaggg gcaacagttg 4920  
ttgcaacctt cgggggcagt aggacaccag cttccgaggc cactcagggg agcagaagag 4980  
agggctgctg gccagcatga cgtaggagaa cataccaatg ctgaaaagct gggaaattcat 5040  
gcagtggaag taggacacaa agaacaggcc aatggatctt gagacatcaa aaaagagcag 5100  
gaaaccagct gagaggtcaa gcagcagccc acccagtgcc acgaccagca ggctagtccag 5160  
ctcctcagac aacagcagtt tgaagggact gaagagccag tgccgggaca aatattccat 5220  
ggaatagcct tcaaccagct ctgcatccag ctttttcaca cccgcaatga agtacacaat 5280  
gaagatctgg ccacggagca ctgcatagtt ccaaaggggc acgtgggcat tcctcctatg 5340  
ggcattcagc agaccgtcca cagaccagta gtggtttgca tccatgaatg ttagctgaaa 5400  
ggccaacaac ccatacagat aggagtgggt gttccatgat gtcttgcca ggagaaacac 5460  
ataccagtat ggcagcagga ataacacaca gcttatccgg tagcacaggc ccagcatcat 5520  
gccagtgcc ccagaaaca tgatggtgta gacaagatac atccagtcaa gtggcagtgg 5580  
gcgtagggca tccagcaagg ggaagcggca cacatccagc ccatcaaggt atttccggtc 5640  
cagagagctg agccccgct cctggggaat gtctagcacc atcaagaacc caaaaagaaa 5700  
acgaaagaca gctaagcttg cagggtcctg ttggtcgattc agcagggcca ccagcctccg 5760  
ccaactggac aaatctgtcc actcaaaacc caagagtctt cctattcggc tgtcctgcct 5820  
gggccctgag atcagttcag ccttgtcttt ctgtacttta tctgagctgg gcgaggtccg 5880

cgcggaacccg gcagacaccg ccattgctct gcggaaggcg caattccacc aactggact 5940  
 agtggatccg agctcggtac caagcttaag tttaaacgct agccagcttg ggtctcccta 6000  
 tagtgagtcg tattaatttc gataagccag taagcagtgg gttctctagt tagccagaga 6060  
 gctttttgca aaagcctagg cctccaaaaa agcctcctca ctacttctgg aatagctcag 6120  
 aggccgaggc ggctcggcc tctgcataaa taaaaaaaaat tagtcagcca tggggcggag 6180  
 aatgggcgga actgggcgga gttagggcg ggatggcg agttagggc gggactatgg 6240  
 ttgctgacta attgagatgc atgctttgca tacttctgcc tgctggggag cctggggact 6300  
 ttccacacc taactgacac acattccaca gccggatcga tgtgggccat cgccctgata 6360  
 gacggttttt cgcccttga cgttgagtc cacgttcttt aatagtggac tctgttcca 6420  
 aactggaaca aactcaacc ctatctcggc ctattctttt gatttataag ggattttgac 6480  
 gatttcggcc tattggtaa aaaatgagct gatttaaca aaatttaacg cgaattaatt 6540  
 ctgtggaatg tgtgtcagtt aggtgtgga aagtccccag gctccccagc aggcagaagt 6600  
 atgcaaagca tgcattcaa ttagtcagca accaggtgtg gaaagtcccc aggcctccca 6660  
 gcaggcagaa gtatgcaaag catgcatctc aattagtcag caaccatagt cccgcccta 6720  
 actccgcca tcccgccct aactccgcc agttccgcc attctccgcc ccatggctga 6780  
 ctaatttttt ttatttatgc agaggccgag gccgctctg cctctgagct attccagaag 6840  
 tagtgaggag gcttttttg aggcctaggc ttttgcaaaa agctcccggg agcttgata 6900  
 tccattttcg gatctgatca agagacagga tgaggatcgt ttgcatgat tgaacaagat 6960  
 ggattgcacg caggttctcc ggccgcttg gtggagagc tattcggcta tgactgggca 7020  
 caacagaaa tcggctgctc tgatgccgcc gtgtccggc tgtcagcga ggggcgccc 7080  
 gttcttttg tcaagaccga cctgtccggc gccctgaatg aactgcagga cgaggcagc 7140  
 cggctatcgt ggctggccac gacggcggt ccttgccag ctgtgctcga cgttgctact 7200  
 gaagcggaa gggactggct gctattgggc gaagtgccgg ggcaggatct cctgtcatct 7260  
 caccttgctc ctgccgaaa agtatccatc atggctgatg caatgcggcg gctgcatacg 7320  
 cttgatccgg ctacctgcc attcgaccac caagcgaac atcgatcga gcgagcacgt 7380

actcggatgg aagccggctc tgtcgatcag gatgatctgg acgaagagca tcaggggctc 7440  
 gcgccagccg aactgttcgc caggctcaag gcgcgcatgc ccgacggcga ggatctcgtc 7500  
 gtgacccatg gcgatgcctg cttgccgaat atcatggtgg aaaatggccg cttttctgga 7560  
 ttcacgcact gtggccggct ggggtgtggcg gaccgctatc aggacatagc gttggctacc 7620  
 cgtgatattg ctgaagagct tggcggcgaa tgggctgacc gcttcctcgt gctttacggt 7680  
 atcgccgctc ccgattcgea gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga 7740  
 gcgggactct ggggttcgaa atgaccgacc aagcgcgccc caacctgcca tcacgagatt 7800  
 tcgattccac cgcgccttc tatgaaaggt tgggcttcgg aatcgtttcc cgggacgccc 7860  
 gctggatgat cctccagcgc ggggatctca tgctggagtt cttcgccac cccaacttgt 7920  
 ttattgcagc ttataatggt tacaataaa gcaatagcat cacaaatttc acaaataaag 7980  
 catttttttc actgcattct agttgtggtt tgtccaaact catcaatgta tcttatcatg 8040  
 tctgtatacc gtcgacctct agctagagct tggcgtaatc atggtcatag ctgtttctctg 8100  
 tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgta 8160  
 aagcctgggg tgccaatga gtgagctaac tcacattaat tgcgttgccg tcaactgcccg 8220  
 ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga 8280  
 gaggcggttt gcgtattggg cgctcttccg cttcctcgtt cactgactcg ctgcgctcgg 8340  
 tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 8400  
 aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 8460  
 gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca 8520  
 aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 8580  
 ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgcccgtt accggatacc 8640  
 tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 8700  
 tcagttcggg ttaggtcgtt cgctccaagc tgggctgtgt gcacgaacct cccgttcagc 8760  
 ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacctggta agacacgact 8820  
 tategccact ggcagcagcc actggtaaca ggattagcag agcagaggtat gtaggcggtg 8880  
 ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg 8940

tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca 9000  
aacaaccac cgctggtagc ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 9060  
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa 9120  
actcacgta agggatthttg gtcacgagat tatcaaaaag gatcttcacc tagatccttt 9180  
taaattaaaa atgaagthttt aaatcaatct aaagtatata tgagtaaact tggctctgaca 9240  
gttaccatg cttaatcagt gaggcaccta tctcagcgat ctgtctattht cgttcatcca 9300  
tagttgctg actccccgtc gtgtagataa ctacgatacg ggagggtta ccatctggcc 9360  
ccagtgtgc aatgataccg cgagaccac gctcaccggc tccagattta tcagcaataa 9420  
accagccagc cggaagggcc gagcgcagaa gtggctctgc aactttatcc gcctccatcc 9480  
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca 9540  
acgttgthtc cattgctaca ggcatcgtgg tgtcacgctc gtcgthttgg atggcttcat 9600  
tcagctccgg ttcccaacga tcaaggcgag ttacatgac cccatgttg tgcaaaaaag 9660  
cggthtagctc ctccggctct ccgatcgttg tcagaagtaa gthggccgca gthttatcac 9720  
tcatgthtat ggcagcactg cataattctc thactgtcat gccatccgta agatgcttht 9780  
ctgtgactgg tgagtactca accaagthcat tctgagaata gthgatgctg cgaccgagth 9840  
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact thaaaagthc 9900  
tcatcatthg aaaaagthct tcggggcgaa aactctcaag gatctthaccg ctgthtgagat 9960  
ccagthcgat gthaccact cthgcaccca actgactthc agcatcttht actthcacca 10020  
gctththctg gthgagcaaaa acaggaaggg aaaaagctgc aaaaagggg .ataagggcga 10080  
cacggaaaatg thgaactc atactctthc thththcaata thattgaagc atthtaccg 10140  
gthattgtct catgagcgga tacataththg aatgththta gaaaaataaa caaatagggg 10200  
thccgcgac atthccccga aaagthccac ctgacgthc 10238

- <210> 14
- <211> 9683
- <212> DNA
- <213> Artificial
- <220>

5



&lt;223&gt; Vector

&lt;400&gt; 14

gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg	60
ccgcatagtt aagccagtat ctgctccctg cttgtgtggt ggaggtcgct gagtagtgcg	120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc	180
ttaggggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt	240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagtcat agcccatata	300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc	360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc	420
attgacgtca atgggtggag tatttacggg aaactgcccc cttggcagta catcaagtgt	480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt	540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca	600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg	660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc	720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg	780
gtaggcgtgt acgggtgggag gtctatataa gcagagctct ctggctaact agagaaccca	840
ctgcttactg gcttatcgaa attaatacga ctcaactatag ggagacccaa gctggctagc	900
gtttaaactt aagcttggtg ccgagctcgg atccactagt ccagtgtggt ggaattgccc	960
ttattcctca gtgaccagg agctgacaca cttagaaggg caattctgca gatcaccatg	1020
cagcgcgtga acatgatcat ggcagaatca ccaggcctca tcaccatctg ccttttagga	1080
tatctactca gtgctgaatg tacagttttt cttgatcatg aaaacgccaa caaaattctg	1140
aatcggccaa agaggtataa ttcaggtaaa ttggaagagt ttgttcaagg gaaccttgag	1200
agagaatgta tggaagaaaa gtgtagtttt gaagaagcac gagaagtttt tgaaaacact	1260
gaaagaacaa ctgaattttg gaagcagtat gttgatggag atcagtgtga gtccaatcca	1320
tgtttaaagt gcggcagttg caaggatgac attaattcct atgaatggtg gtgtcccttt	1380
ggatttgaag gaaagaactg tgaattagat gtaacatgta acattaagaa tggcagatgc	1440

gagcagtttt gtaaaaatag tgctgataac aagggtggttt gctcctgtac tgagggatat 1500  
cgacttgcag aaaaccagaa gtctgtgtaa ccagcagtcg catttccatg tggaagagtt 1560  
tctgtttcac aaacttctaa gctcaccctg gctgagactg tttttcctga tgtggactat 1620  
gtaaattcta ctgaagctga aaccattttg gataacatca ctcaaagcac ccaatcattt 1680  
aatgacttca ctcggttgtg tgggtggagaa gatgccaaac cagggtcaatt cccttggcag 1740  
gttgttttga atggtaaagt tgatgcattc tgtggaggct ctatcgtaa tgaaaaatgg 1800  
attgtaactg ctgcccactg tgttgaaact ggtgttaaaa ttacagttgt cgcaggtgaa 1860  
cataatattg aggagacaga acatacagag caaaagcgaa atgtgattcg aattattcct 1920  
caccacaact acaatgcagc tattaataag tacaaccatg acattgccct tctggaactg 1980  
gacgaacct tagtgctaaa cagctacgtt acacctattt gcattgctga caaggaatac 2040  
acgaacatct tctcaaatt tggatctggc tatgtaagtg gctggggaag ggtcttccac 2100  
aaagggagat cagctttagt tcttcagtac cttagagttc cacttgttga ccgagccaca 2160  
tgtcttcgat ctacaaagtt caccatctat aacaacatgt tctgtgctgg cttccatgaa 2220  
ggaggtagag attcatgtca aggagatagt gggggacccc atgttactga agtgggaaggg 2280  
accagtttct taactggaat tattagctgg ggtgaagagt gtgcaatgaa aggcaaatat 2340  
ggaatatata ccaaggtatc ccggtatgtc aactggatta aggaaaaaac aaagctcact 2400  
taatgaaaga tggatttcca aggatccagc acagtggcgg ccgctcgagt ctagagggcc 2460  
cgtttaaacc cgctgatcag cctcgactgt gccttctagt tgccagccat ctgttgtttg 2520  
ccccccccc gtgccttctt tgaccctgga aggtgccact cccactgtcc tttcctaata 2580  
aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg ggggtggggt 2640  
ggggcaggac agcaaggggg aggattggga agacaatagc aggcatgctg gggatgctgt 2700  
gggctctatg gcttctgagg cggaaagaac cagctggggc tctaggggggt atccccacgc 2760  
gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac 2820  
acttgcagc gccctagcgc ccgctcctt cgctttcttc ccttccttcc tcgccaagtt 2880  
cgccggcttt ccccgtaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc 2940

ES 2 357 603 T3

ttacggcac cttcgacccc aaaaaacttg attagggctg tggaatgtgt gtcagttagg 3000  
 gtgtggaaag tccccaggct ccccagcagg cagaagtatg caaagcatgc atctcaatta 3060  
 gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta tgcaaagcat 3120  
 gcatctcaat tagtcagcaa ccatagtccc gccccctaact ccgcccattcc cgcccctaac 3180  
 tccgcccagt tccgcccatt ctccgcccga ttggtgacta atttttttta tttatgcaga 3240  
 ggccgaggcc gcctcggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg 3300  
 cctaggcttt tgcaaaaagc tctctggcta actagagaac ccactgctta ctggcttacc 3360  
 gaaattaata cgactcacta tagggagacc caagctggct agcgtttaa cttagcttg 3420  
 gtaccgagct cggatccact agtccagtgt ggtggaattg cccttccgc agagcaatgg 3480  
 cgggtgtctgc cgggtccgcy cggacctcgc ccagctcaga taaagtacag aaagacaagg 3540  
 ctgaactgat ctcagggccc aggcaggaca gccgaatagg gaaactcttg ggttttgagt 3600  
 ggacagattt gtccagttgg cggaggctgg tgaccctget gaatcgacca acggaccctg 3660  
 caagcttagc tgtctttcgt tttcttttg ggttcttgat ggtgctagac attccccagg 3720  
 agcgggggct cagctctctg gaccgaaat acctgatgg gctggatgtg tgccgcttcc 3780  
 ccttgctgga tgccctacgc ccactgccac ttgactggat gtatcttctc tacaccatca 3840  
 tgtttctggg ggcaactggc atgatgctgg gcctgtgcta ccggataagc tgtgtgttat 3900  
 tctgtctgcc atactggtat gtgtttctcc tggacaagac atcatggaac aaccactcct 3960  
 atctgtatgg gttgttggcc tttcagctaa cattcatgga tgcaaaccac tactggctctg 4020  
 tggacggtct gctgaatgcc cataggagga atgcccacgt gccctttgg aactatgcag 4080  
 tgctccgtgg ccagatcttc attgtgtact tcattgctgg tgtgaaaaag ctggatgcag 4140  
 actgggttga agctattcc atggaatatt tgtcccgca ctggctcttc agtcccttca 4200  
 aactgctggt gtctgaggag ctgactagcc tgctggctgt gcaactgggt gggctgctgc 4260  
 ttgacctctc agctggtttc ctgctctttt ttgatgtctc aagatccatt ggctgttct 4320  
 ttgtgtccta ctccactgc atgaattccc agcttttcag cattggtatg ttctcctacg 4380  
 tcatgctggc cagcagccct ctcttctgct ccctgagtg gcctcggag ctggtgtcct 4440  
 actgcccccg aaggttgcaa caactgttgc ccctcaaggc agcccctcag cccagtgttt 4500

ES 2 357 603 T3

cctgtgtgta taagaggagc cggggcaaaa gtggccagaa gccagggctg cgccatcagc 4560  
 tgggagctgc cttcacctg ctctacctcc tggagcagct attcctgccc tattctcatt 4620  
 ttctcaccca gggctataac aactggacaa atgggctgta tggctattcc tgggacatga 4680  
 tgggtgactc ccgctcccac cagcacgtga agatcaccta ccgtgatggc cgactggcg 4740  
 aactgggcta ccttaaccct ggggtattta cacagagtcg gcgatggaag gatcatgcag 4800  
 acatgctgaa gcaatatgcc acttgectga gccgctgct tcccaagtat aatgtcactg 4860  
 agccccagat ctactttgat atttgggtct ccatcaatga ccgcttcag cagaggattt 4920  
 ttgacctcg tgtggacatc gtgcaggccg cttggtcacc cttcagcgc acatcctggg 4980  
 tgcaaccact cttgatggac ctgtctccct ggagggccaa gttacaggaa atcaagagca 5040  
 gcctagacaa ccacactgag gtggcttca ttgcagattt ccctggactg cacttggaga 5100  
 attttgtgag tgaagacctg ggcaacacta gcattccagct gctgcagggg gaagtgactg 5160  
 tggagcttgt ggcagaacag aagaaccaga ctcttcgaga gggagaaaa atgcagttgc 5220  
 ctgctggtga gtaccataag gtgtatacga catcacctag cccttcttgc tacatgtacg 5280  
 tctatgtcaa cactacagag cttgactgg agcaagacct ggcatatctg caagaattaa 5340  
 aggaaaagg tggagaatgga agtgaaacag ggcctctacc ccagagctg cagcctctgt 5400  
 tgggaagggga agtaaaagg ggcctgagc caacacctct gggtcagacc tttcttagac 5460  
 gccaaacaag gctccaggag attgaacgcc ggcgaaatac tcctttccat gagegattct 5520  
 tccgcttctt gttgcgaaag ctctatgtct ttgcgccag cttcctgatg acttgtatct 5580  
 cacttcgaaa tctgatatta ggccgtcctt ccctggagca gctggcccag gaggtgactt 5640  
 atgcaaactt gagacccttt gaggcagttg gagaactgaa tcctcaaac acggattctt 5700  
 cacattctaa tcctcctgag tcaaatcctg atcctgtcca ctcaagttc tgaagggggc 5760  
 cagatgttgg aagggcaatt cgagtctaga gggcccggcc tgatagacgg ttttctgccc 5820  
 tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact 5880  
 caacctatc tcggtctatt cttttgattt ataaggatt ttgccgattt cggcctattg 5940  
 gttaaaaaat gagctgattt aacaaaaatt taacgcgaat taattctgtg gaatgtgtgt 6000

cagttagggg gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 6060  
ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg 6120  
caaagcatgc atctcaatta gtcagcaacc atagtcccgc ccctaactcc gcccateccg 6180  
cccctaactc cgcccagttc cgcccattct ccgcccctg gctgactaat tttttttatt 6240  
tatgcagagg ccgaggccgc ctctgectct gagctattcc agaagtagtg aggaggcttt 6300  
tttggaggcc taggcttttg caaaaagctc ccgggagctt gtatatccat tttcggatct 6360  
gatcaagaga caggatgagg atcgtttcgc atgattgaac aagatggatt gcacgcagg 6420  
tctccggccg ctgggtgga gaggctattc ggctatgact gggcacaaca gacaatcggc 6480  
tgctctgatg ccgccgtgtt ccggctgtca gcgcaggggc gcccggttct ttttgcaag 6540  
accgacctgt ccggtgccct gaatgaactg caggacgagg cagcgcggct atcgtggctg 6600  
gccacgacgg gcgttccttg cgcagctgtg ctcgacgttg tcaactgaagc gggaagggac 6660  
tggctgctat tgggcgaagt gccggggcag gatctcctgt catctcacct tgctcctgcc 6720  
gagaaagtat ccatcatggc tgatgcaatg cggcggetgc atacgcttga tccggctacc 6780  
tgcccattcg accaccaagc gaaacatcgc atcgagcgag cacgtactcg gatggaagcc 6840  
ggctctgtcg atcaggatga tctggacgaa gagcatcagg ggctcgcgcc agccgaactg 6900  
ttcgccaggc tcaaggecgc catgcccgc ggcgaggatc tcgtcgtgac ccatggcgat 6960  
gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt ctggattcat cgactgtggc 7020  
cggctgggtg tggcggaccg ctatcaggac atagcgttgg ctaccctga tattgctgaa 7080  
gagcttggcg gcgaatgggc tgaccgcttc ctctgtcttt acggtatcgc cgctcccgat 7140  
tcgcagcgca tcgcctteta tcgccttctt gacgagttct tctgagcggg actctgggg 7200  
tcgaaatgac cgaccaagcg acgcccacc tgccatcacg agatttcgat tccaccgccg 7260  
ccttctatga aaggttgggc ttcggaatcg ttttccggga cgcgggctgg atgatcctcc 7320  
agcgcgggga tctcatgctg gagttctctg cccaccccaa cttgtttatt gcagcttata 7380  
atggttacia ataaagcaat agcatcacia atttcacaaa taaagcattt ttttactgc 7440  
attctagttg tggtttgcc aaactcatca atgtatctta tcatgtctgt ataccgtcga 7500  
cctctagcta gagcttggcg taatcatggt catagctggt tcctgtgtga aattgttatc 7560

ES 2 357 603 T3

cgctcacaat tccacacaac atacgagcgg gaagcataaa gtgtaaagcc tgggggtgcct 7620  
 aatgagtgag ctaactcaca ttaattgcgt tgcgctcact gcccgcttcc cagtcgggaa 7680  
 acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggg ggtttgcgta 7740  
 ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctgggtcgtt cggctgcggc 7800  
 gagcggatc agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg 7860  
 caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt 7920  
 tgctggcgtt tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa 7980  
 gtcagagggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct 8040  
 cectcgtgcg ctctcctggt cegaccctgc cgcttaccgg atacctgtcc gcctttctcc 8100  
 cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcgggtgtagg 8160  
 tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccagc cgctgcgcct 8220  
 tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccaactggcag 8280  
 cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca gaggctctga 8340  
 agtgggtggc taactacggc tacactagaa gaacagtatt tggtatctgc gctctgctga 8400  
 agccagttac cttcggaaaa agagtggta gctcttgatc cggcaaaaa accaccgctg 8460  
 gtagcggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag 8520  
 atcctttgat cttttctacg gggctcgcg ctgagtgga cgaaaactca cgtaagggg 8580  
 ttttggcat gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa 8640  
 gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac caatgcttaa 8700  
 tcagtgaggc acctatctca gcgatctgc tatttcgctc atccatagtt gcctgactcc 8760  
 ccgctcgtgta gataactacg atacgggagg gcttaccatc tggcccagc gctgcaatga 8820  
 taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa 8880  
 gggccgagcg cagaagtggc cctgcaactt tatccgctc catccagctc attaatggt 8940  
 gccgggaagc tagagtaagt agttcgcag ttaatagttt gcgcaacgtt gttgccattg 9000  
 ctacaggeat cgtgggtgca cgctcgtcgt ttggatggc ttcattcagc tccggttccc 9060

aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcgggt agtccttcg 9120  
 gtctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 9180  
 cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt 9240  
 actcaaccaa gtcattctga gaatagtgtg tgcggcgacc gagttgctct tgccccggct 9300  
 caatacggga taataccgcy ccacatagca gaactttaa agtgetcatc attggaaaac 9360  
 gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac 9420  
 ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag 9480  
 caaaaacagg aaggcaaat gccgcaaaa agggaataag ggcgacacgg aatgttgaa 9540  
 tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 9600  
 ggggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 9660  
 cccgaaaagt gccacctgac gtc 9683

- <210> 15
- <211> 9781
- <212> DNA
- 5 <213> Artificial
- <220>
- <223> Vector
- <400> 15

gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60  
 ccgcatagtt aagccagtat ctgctccctg cttgtgtggt ggaggctcgt gagtagtgcg 120  
 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180  
 ttagggtag gcgttttgcy ctgcttcgcy atgtacgggc cagatatacg cgttgacatt 240  
 gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300  
 tggagttccg cgttacataa cttacggtaa atggcccgc tgctgaccg cccaacgacc 360  
 cccgccatt gacgtcaata atgacgtatg ttccatagt aacgccaata gggactttcc 420  
 attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 480  
 atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540  
 atgcccagta catgacctta tgggactttc ctacttgcca gtacatctac gtattagtca 600

tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg	660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc	720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg	780
gtaggcgtgt acgggtgggag gtctatataa gcagagctct ctggctaact agagaacca	840
ctgcttactg gcttatcgaa attaatacga ctcaactatag ggagacccaa gctggctagc	900
gtttaaactt aagcttggtg ccgagctcgg atccactagt ccagtgtggt ggaattgccc	960
ttattcctca gtgaccagg agctgacaca cttagaaggg caattctgca gataccatgc	1020
agcgcgtgaa catgatcatg gcagaatcac caggcctcat caccatctgc cttttaggat	1080
atctactcag tgctgaatgt acagttttc ttgatcatga aaacgccaac aaaattctga	1140
atcggccaaa gaggtataat tcaggtaaat tggaagagtt tgttcaaggg aaccttgaga	1200
gagaatgtat ggaagaaaag tgtagttttg aagaagcacg agaagttttt gaaaacactg	1260
aaagaacaac tgaattttgg aagcagtatg ttgatggaga tcagtgtgag tccaatccat	1320
gtttaaatgg cggcagttgc aaggatgaca ttaattccta tgaatggttg tgtccctttg	1380
gatttgaagg aaagaactgt gaattagatg taacatgtaa cattaagaat ggcagatgcg	1440
agcagttttg taaaaatagt gctgataaca aggtggtttg ctctgtact gagggatac	1500
gacttgcaga aaaccagaag tctgtgaac cagcagtgcc atttccatgt ggaagagttt	1560
ctgtttcaca aacttctaag ctcaccctg ctgagactgt ttttctgat gtggactatg	1620
taaattctac tgaagctgaa accattttgg ataacatcac tcaaagcacc caatcattta	1680
atgacttcac tcgggttgtt ggtggagaag atgccaaacc aggtcaattc ccttggcagg	1740
ttgttttgaa tggtaaagtt gatgcattct gtggaggctc tatcgttaat gaaaaatgga	1800
ttgtaactgc tgcccactgt gttgaaactg gtgttaaaat tacagttgtc gcaggtgaac	1860
ataatattga ggagacagaa catacagagc aaaagcgaaa tgtgattcga attattcctc	1920
accacaacta caatgcagct attaataagt acaaccatga cattgccctt ctggaactgg	1980
acgaaccctt agtgctaaac agctacgtta cacctatttg cattgctgac aaggaataca	2040
cgaacatctt cctcaaattt ggatctggct atgtaagtgg ctggggaagg gtcttcaca	2100



aaggagatc agcttagt cttcagtacc ttagagttcc acttggtgac cgagccacat 2160  
gtcttcgatc tacaaagttc accatctata acaacatggt ctgtgctggc ttccatgaag 2220  
gaggtagaga ttcatgtcaa ggagatagtg ggggaccca tgttactgaa gtggaagga 2280  
ccagtttctt aactggaatt attagctggg gtgaagagtg tgcaatgaaa ggcaaatatg 2340  
gaatatatac caaggtatcc cggtatgtca actggattaa ggaaaaaca aagctcactt 2400  
aatgaaagat ggatttccaa ggatccagca cagtggcggc cgctcggttc ctagagggcc 2460  
cgtttaaacc cgctgatcag cctcgactgt gccttctagt tgccagccat ctggtggttg 2520  
cccccccc gtgccttctt tgaccctgga aggtgccact cccactgtcc ttctctaata 2580  
aatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg ggggtgggt 2640  
ggggcaggac agcaaggggg aggattggga agacaatagc aggcattgctg gggatgctgt 2700  
gggctctatg gcttctgagg cggaaagaac cagctggggc tctaggggtt atccccacgc 2760  
gccctgtagc ggcgcatcaa gcgcgggggg tgtggtggtt acgcgagcg tgaccgctac 2820  
acttgccagc gccctagcgc ccgctcctt cgtttcttc ccttccttc tcgccacgtt 2880  
cgccggcttt ccccgtaag ctctaaatcg ggggctcct ttagggttcc gatttagtgc 2940  
tttacggcac ctcgaccca aaaaactga ttaggggtgat ggttcacatc gatgcaattt 3000  
cctcatttta ttaggaaagg acagtgggag tggcacctc cagggtcaag gaaggcacgg 3060  
gggaggggca aacaacagat ggctggcaac tagaaggcac agtcgaggct gatcagcggg 3120  
tttaaacggg ccctctagac tcgaattgcc cttccaacat ctggccccct tcagaactct 3180  
gagtggacag gatcaggatt tgactcagga ggattagaat gtgaagaatc cgtgtttgag 3240  
ggattcagtt ctccaactgc ctcaaagggt ctcaagttg cataagtac ctctgggcc 3300  
agctgctcca ggaaggacg gcctaatac agatttcgaa gtgagataca agtcatcagg 3360  
aagctgcggc gaaagacata gagctttcgc aacaagaagc ggaagaatcg ctcatggaaa 3420  
ggagtatttc gccggcgttc aatctctgg agcctttggt ggcgtctaag aaaggtctga 3480  
accagaggtg ttggctcagg gccccctttt acttccccct ccaacagagg ctgcagctct 3540  
gggggtagag gccctgttcc acttccatcc tccacctttt cctttaatte ttgcagatat 3600  
gccaggtctt gctccagtgc aagctctgta gtgttgacat agacgtacat gtagcaagaa 3660

gggctaggtg atgtcgtata caccttatgg tactcaccag caggcaactg cattttttct 3720  
 ccctctcgaa gaggctgggt cttctgttct gccacaagct ccacagtcac ttccccctgc 3780  
 agcagctgga tgctagtgtt gcccagggtct tcaactcacia aattctccaa gtgcagtcca 3840  
 gggaaatctg caatgaagac cacctcagtg tggttgtcta ggctgctctt gatttctctg 3900  
 aacttggccc tccagggaga caggtccatc aagagtgggt gcaccaggga tgtgctgctga 3960  
 aagggtgacc aagcggcctg cacgatgtcc acacgagggg caaaaatcct ctgctggaag 4020  
 cggtcattga tggagaccca aatatcaaag tagatctggg gctcagtgac attatacttg 4080  
 ggaagcaggc ggctcaggca agtggcatat tgcttcagca tgtctgcatg atccttccat 4140  
 cgccgactct gtgtaaatac cccaggggta aggtagccca gttcgccagt ggggccatca 4200  
 cggtaggtga tcttcacgtg ctgggtgggag cgggagtgca ccatcatgtc ccaggaatag 4260  
 ccatacagcc catttgtcca gttgttatag ccctgggtga gaaaatgaga atagggcagg 4320  
 aatagctgct ccaggaggta gagcagggtg aaggcagctc ccagctgatg ggcagccct 4380  
 ggcttctggc cacttttgcc ccggctcctc ttatacacac aggaaact gggctgaggg 4440  
 gctgccttga ggggcaacag ttgttgcaac cttcgggggc agtaggacac cagcttccga 4500  
 ggccactcag gggagcagaa gagagggctg ctggccagca tgacgtagga gaacatacca 4560  
 atgctgaaaa gctgggaatt catgcagtg aagtaggaca caaagaacag gccaatggat 4620  
 cttgagacat caaaaaagag caggaaacca gctgagaggt caagcagcag cccaccccag 4680  
 tgcacgacca gcaggctagt cagctcctca gacaacagca gtttgaaggg actgaagagc 4740  
 cagtgccggg acaaatatc catggaatag cctcaacce agtctgcatc cagctttttc 4800  
 acaccgcaa tgaagtacac aatgaagatc tggccaagga gactgcata gttccaaagg 4860  
 ggcacgtggg cattcctcct atgggcatc agcagaccgt ccacagacca gtagtggttt 4920  
 gcatccatga atgttagctg aaaggccaac aaccataca gataggagtg gttgttccat 4980  
 gatgtcttgt ccaggagaaa cacataccag tatggcagca ggaataacac acagcttatc 5040  
 cggtagcaca ggcccagcat catgcccagt gccccagaa acatgatggt gtagacaaga 5100  
 tacatccagt caagtggcag tgggcgtagg gcatccagca aggggaagcg gcacacatcc 5160

agcccatcaa ggtatttccg gtccagagag ctgagccccc gtccttgggg aatgtctagc 5220  
 accatcaaga acccaaaaag aaaacgaaag acagctaagc ttgcaggggc cgttggtcga 5280  
 ttcagcaggg tcaccagcct ccgccaactg gacaaatctg tccactcaaa acccaagagt 5340  
 ttccctattc gggtgtcctg cctgggcoct gagatcagtt cagccttgc tttctgtact 5400  
 ttatctgagc tgggcgaggt ccgcgcgagc ccggcagaca ccgccattgc tctgcgga 5460  
 gggcaattcc accacactgg actagtggat ccgagctcgg taccaagcct aagttaaac 5520  
 gctagccagc ttgggtctcc ctatagttag tcgtattaat ttcgataagc cagtaagcag 5580  
 tgggttctct agttagccag agagcttttt gcaaaagcct aggcctcaa aaaagcctcc 5640  
 tcaactactc tggaatagct cagaggccga ggcggcctcg gcctctgcat aaataaaaa 5700  
 aattagtcag ccatggggcg gagaatgggc ggaactgggc ggagttaggg gcgggatggg 5760  
 cggagttagg ggcgggacta tggttgctga ctaattgaga tgcatgcttt gcatactct 5820  
 gcctgctggg gagcctgggg actttccaca ccctaactga cacacattcc acagccggat 5880  
 cgatgtgggc catcgcctg atagacgggt tttcgcctt tgacgttga gtccacgttc 5940  
 ttaaatagtg gactcttgtt ccaaactgga acaacactca accctatctc ggtctattct 6000  
 tttgatttat aagggatttt gccgatttcg gcctattggt taaaaaatga gctgatttaa 6060  
 caaaaattta acgcaatta attctgtgga atgtgtgtca gttaggggtg ggaaagtccc 6120  
 caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca gcaaccagg 6180  
 gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat ctcaattagt 6240  
 cagcaaccat agtcccgcct ctaactccgc ccatcccgc cctaactccg cccagttccg 6300  
 cccattctcc gcccattggc tgactaattt tttttattta tgcagaggcc gaggccgcct 6360  
 ctgcctctga gctattccag aagtagtgag gaggcttttt tggaggccta ggcttttgca 6420  
 aaaagctccc gggagcttgt atatccattt tcggatctga tcaagagaca ggatgaggat 6480  
 cgtttcgcct gattgaacaa gatggattgc acgcaggttc tccggccgct tgggtggaga 6540  
 ggctattcgg ctatgactgg gcacaacaga caatcggctg ctctgatgcc gccgtgttcc 6600  
 ggctgtcagc gcagggggcg ccggttcttt ttgtcaagac cgacctgtcc ggtgccctga 6660  
 atgaactgca ggacgaggca gcgcggtctat cgtggctggc cagcaggggc gttccttgcg 6720

cagctgtgct cgacgttgtc actgaagcgg gaagggactg gctgctattg ggccaagtgc 6780  
cggggcagga tctcctgtca tctcaccttg ctctgcccga gaaagtatcc atcatggctg 6840  
atgcaatgcg gcggctgcat acgcttgatc cggctacctg cccattcgac caccaagcga 6900  
aacatcgcat cgagcgagca cgtactcgga tgggaagccgg tcttgctgat caggatgatc 6960  
tggacgaaga gcatcagggg ctgcgcgccag ccgaactggt cgccaggctc aaggcgcgca 7020  
tgcccgacgg cgaggatctc gtcgtgacct atggcgatgc ctgcttgccg aatatcatgg 7080  
tggaaaatgg cgcttttct ggattcatcg actgtggccg gctgggtgtg gcggaccgct 7140  
atcaggacat agcgttggct acccgtgata ttgctgaaga gcttggcggc gaatgggctg 7200  
accgcttctc cgtgctttac ggtatcgccg ctcccgatc gcagcgcac gccttctatc 7260  
gccttcttga cgagttcttc tgagcggggac tctgggggttc gaaatgaccg accaagcgac 7320  
gcccAACctg ccatcacgag atttogatc caccgcgcc ttctatgaaa ggttgggctt 7380  
cggaatcgtt tccggggacg ccggctggat gatcctccag cgcggggatc tcatgctgga 7440  
gttcttgcgc caccCCAact tgtttattgc agcttataat ggttacaaat aaagcaatag 7500  
catcacaAat ttcacaaata aagcattttt ttcactgcat tctagttgtg gtttgtccaa 7560  
actcatcaat gtatcttata atgtctgtat accgtcgacc tctagctaga gcttggcgta 7620  
atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc cacacaacat 7680  
acgagccgga agcataaagt gtaaagcctg gggtgccaa tgagtgagct aactcacatt 7740  
aattgcgttg cgtcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta 7800  
atgaatcggc caacgcgagg ggagaggcgg tttgcgtatt gggcgctctt ccgcttctc 7860  
gctcactgac tcgctgcgct cggctgttcg gctgcccga gcggtatcag ctcaactcaa 7920  
ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 7980  
agccagcaa aaggccagga accgtaaaaa ggccgcgctg ctggcgtttt tccataggct 8040  
ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 8100  
aggactataa agataccagg cgtttcccc tggaagctcc ctctgctgct ctctgttcc 8160  
gaccctgccg cttaccggat acctgtccgc ctttctcct tcgggaagcg tggcgcttcc 8220

ES 2 357 603 T3

tcatagctca	cgctgtaggt	atctcagttc	gggtgtaggtc	gttcgctcca	agctgggctg	8280
tgtgcacgaa	cccccgttc	agccccgaccg	ctgcgccctta	tccggtaact	atcgtcttga	8340
gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	gccactggta	acaggattag	8400
cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	tggtggccta	actacggcta	8460
cactagaaga	acagtatttg	gtatctgcg	tctgctgaag	ccagttacct	tccgaaaaag	8520
agttggtagc	tcttgatccg	gcaaacaaaac	caccgctgggt	agcgggtttt	ttgtttgcaa	8580
gcagcagatt	acgcgcagaa	aaaaaggatc	tcaagaagat	cctttgatct	tttctacggg	8640
gtctgacgct	cagtggaacg	aaaactcacg	ttaagggatt	ttggatcatga	gattatcaaa	8700
aaggatcttc	acctagatcc	ttttaaatta	aaaatgaagt	tttaaatcaa	tctaaagtat	8760
atatgagtaa	acttggctctg	acagttacca	atgcttaatc	agtgaggcac	ctatctcagc	8820
gatctgtcta	tttcgttcat	ccatagttgc	ctgactcccc	gtcgtgtaga	taactacgat	8880
acgggagggc	ttaccatctg	gccccagtgc	tgcaatgata	ccgcgagacc	cacgctcacc	8940
ggctccagat	ttatcagcaa	taaaccagcc	agccggaagg	gccgagcgca	gaagtgggtcc	9000
tgcaacttta	tccgcctcca	tccagtctat	taattgttgc	cgggaagcta	gagtaagtag	9060
ttcggccagtt	aatagtttgc	gcaacgttgt	tgccattgct	acaggcatcg	tggtgtcacg	9120
ctcgtcgttt	ggtatggctt	cattcagctc	cggttcccaa	cgatcaaggc	gagttacatg	9180
atcccccatg	ttgtgcaaaa	aagcggttag	ctccttcgggt	cctccgatcg	ttgtcagaag	9240
taagttggcc	gcagtgttat	cactcatggt	tatggcagca	ctgcataatt	ctcttactgt	9300
catgccatcc	gtaagatgct	tttctgtgac	tggtgagtac	tcaaccaagt	cattctgaga	9360
atagtgtatg	cggcgaccga	gttgcctctg	cccggcggtca	atacgggata	ataccgcgcc	9420
acatagcaga	actttaaaag	tgctcatcat	tggaaaacgt	tcttcggggc	gaaaactctc	9480
aaggatctta	ccgctgttga	gatccagttc	gatgtaaccc	actcgtgcac	ccaactgatc	9540
ttcagcatct	tttactttca	ccagcgtttc	tgggtgagca	aaaacaggaa	ggcaaaatgc	9600
cgcaaaaaag	ggaataaggg	cgacacggaa	atgttgaata	ctcatactct	tcctttttca	9660
atattattga	agcatttatc	agggttattg	tctcatgagc	ggatacatat	ttgaatgtat	9720
ttagaaaaat	aaacaaatag	gggttccgcg	cacatttccc	cgaaaagtgc	cacctgacgt	9780
c						9781

## REIVINDICACIONES

1. Una célula huésped eucariota que comprende al menos un vector de expresión que comprende:  
una molécula de ácido nucleico que codifica una proteína que requiere gamma-carboxilación y secuencias control asociadas a la expresión que comprenden un primer promotor y
- 5 una molécula de ácido nucleico que codifica una  $\gamma$ -glutamil carboxilasa y secuencias control asociadas a la expresión que comprenden un segundo promotor,  
en el que el primer promotor se selecciona de promotor temprano inmediato del citomegalovirus humano (hCMV), pEF-1 $\alpha$ , pRSV o pUbC, y
- 10 el segundo promotor se selecciona de promotor temprano inmediato de SV40, promotor FIX minimizado o promotor de quinasa timidina de HSV.
2. La célula huésped que se reivindica en la reivindicación 1, en la que las moléculas de ácido nucleico están dentro de un vector de expresión sencillo.
3. La célula huésped que se reivindica en la reivindicación 1 o 2, en la que el primer promotor es el promotor temprano inmediato del citomegalovirus humano (hCMV) y el segundo promotor es el promotor temprano de SV40.
- 15 4. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 3, en la que la proteína que requiere gamma-carboxilación se selecciona del grupo que consiste en: factor de coagulación VII, factor de coagulación IX, protrombina (factor de coagulación II), factor de coagulación X, proteína C, proteína S, proteína Z, proteína Bone Gla, proteína de Matriz Gla, proteína específica de detención del crecimiento 6 y proteína del tipo FXa de Acanthophiinae.
- 20 5. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 4, en la que la proteína que requiere gamma-carboxilación es un factor de coagulación dependiente de la vitamina K.
6. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 5, en la que la proteína que requiere gamma-carboxilación es el Factor IX.
- 25 7. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 5, en la que la proteína que requiere gamma-carboxilación es el Factor X.
8. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 5, en la que la proteína que requiere gamma-carboxilación es la protrombina.
- 30 9. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 5, en la que la proteína que requiere gamma-carboxilación es la proteína C.
10. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 5, en la que la proteína que requiere gamma-carboxilación es el Factor VII.
11. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 10, en la que la proteína que requiere gamma-carboxilación es una proteína humana.
- 35 12. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 11, en la que la  $\gamma$ -glutamil carboxilasa es una proteína humana.
13. Una célula como se reivindica en cualquiera de las reivindicaciones 1 a 12, en la que la célula se selecciona del grupo que consiste en: una célula de mamífero, célula de levadura o célula de insecto.
- 40 14. Una célula como se reivindica en la reivindicación 13, en la que la célula es una célula de mamífero.
15. Un vector que comprende una molécula de ácido nucleico que codifica una proteína que requiere gamma-carboxilación y secuencias control asociadas a la expresión que comprenden un primer promotor y una molécula de ácido nucleico que codifica una  $\gamma$ -glutamil carboxilasa y secuencias control asociadas a la expresión que comprenden un segundo promotor,
- 45 en el que el primer promotor se selecciona de promotor temprano inmediato del citomegalovirus humano (hCMV), pEF-1 $\alpha$ , pRSV o pUbC, y  
el segundo promotor se selecciona de promotor temprano inmediato de SV40, promotor FIX minimizado o promotor de quinasa timidina de HSV.
- 50 16. Un vector como se reivindica en la reivindicación 15, en el que el primer promotor es el promotor temprano inmediato del citomegalovirus humano (hCMV) y el segundo promotor es el promotor temprano de SV40.

- 5 17. Un vector como se reivindica en la reivindicación 15 o 16, en el que la proteína que requiere gamma-carboxilación se selecciona del grupo que consiste en: factor de coagulación VII, factor de coagulación IX, protrombina (factor de coagulación II), factor de coagulación X, proteína C, proteína S, proteína Z, proteína Bone Gla, proteína de Matriz Gla, proteína específica de detención del crecimiento 6 y proteína del tipo FXa de Acanthophiinae.
18. Un vector como se reivindica en cualquiera de las reivindicaciones 15 a 17, en el que la proteína que requiere gamma-carboxilación es el Factor IX.
19. Un vector como se reivindica en cualquiera de las reivindicaciones 15 a 17, en el que la proteína que requiere gamma-carboxilación es el Factor X.
- 10 20. Un vector como se reivindica en cualquiera de las reivindicaciones 15 a 17, en el que la proteína que requiere gamma-carboxilación es la protrombina.
21. Un vector como se reivindica en cualquiera de las reivindicaciones 15 a 17, en el que la proteína que requiere gamma-carboxilación es la proteína C.
- 15 22. Un vector como se reivindica en cualquiera de las reivindicaciones 15 a 17, en el que la proteína que requiere gamma-carboxilación es el Factor VII.
23. Un método para producir la proteína gamma-carboxilada que comprende:
- (i) cultivar una célula eucariota que comprende al menos un vector de expresión que comprende una molécula de ácido nucleico que codifica una proteína que requiere gamma-carboxilación y secuencias control asociadas a la expresión que comprenden un primer promotor y
- 20 una molécula de ácido nucleico que codifica una  $\gamma$ -glutamyl carboxilasa y secuencias control asociadas a la expresión que comprenden un segundo promotor,
- en el que el primer promotor se selecciona del promotor temprano inmediato del citomegalovirus humano (hCMV), pEF-1 $\alpha$ , pRSV y pUbC, y
- 25 el segundo promotor se selecciona del promotor temprano inmediato de SV40, promotor FIX minimizado o promotor de quinasa timidina de HSV, y
- (ii) aislar la proteína gamma-carboxilada.
24. Un método como se reivindica en la reivindicación 23, en el que la proteína gamma-carboxilada se selecciona del grupo que consiste en: factor de coagulación VII, factor de coagulación IX, protrombina (factor de coagulación II), factor de coagulación X, proteína C, proteína S, proteína Z, proteína Bone Gla, proteína de Matriz Gla, proteína específica de detención del crecimiento 6 y proteína del tipo FXa de Acanthophiinae.
- 30 25. Un método como se reivindica en la reivindicación 23 o 24, en el que la proteína que requiere gamma-carboxilación es un factor de coagulación dependiente de la vitamina K o su forma activada.
26. Un método como se reivindica en la reivindicación 23 o 25, en el que la proteína que requiere gamma-carboxilación es el Factor IX o su forma activada.
- 35 27. Un método como se reivindica en la reivindicación 23 o 25, en el que la proteína que requiere gamma-carboxilación es el Factor X o su forma activada.
28. Un método como se reivindica en la reivindicación 23 o 25, en el que la proteína que requiere gamma-carboxilación es la protrombina o su forma activada.
- 40 29. Un método como se reivindica en la reivindicación 23 o 25, en el que la proteína que requiere gamma-carboxilación es la proteína C o su forma activada.
30. Un método como se reivindica en la reivindicación 23 o 25, en el que la proteína que requiere gamma-carboxilación es el Factor VII o su forma activada.
31. Un método para producir una composición farmacéutica adecuada para inducir la coagulación de la sangre o promover un incremento o decrecimiento de la coagulación, que comprende la etapas de
- 45 (i) expresar una proteína carboxilada activa en una célula huésped de acuerdo con la reivindicación 1,
- (ii) purificar la proteína carboxilada activa, y
- (iii) mezclar la proteína carboxilada purificada con uno o varios vehículos farmacéuticamente aceptables o excipientes.

Figura 1a

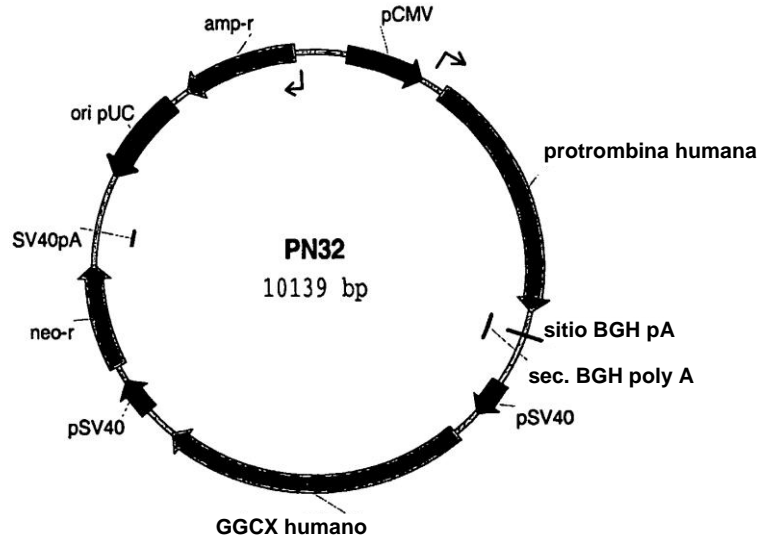


Figura 1b

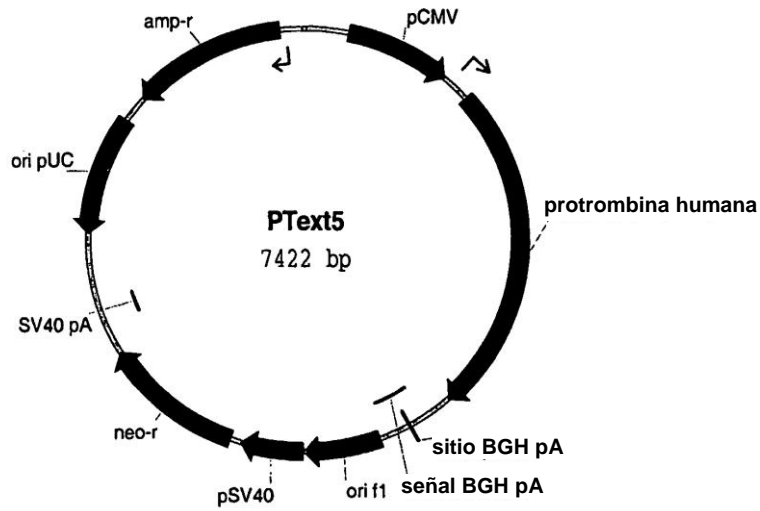




Figura 2

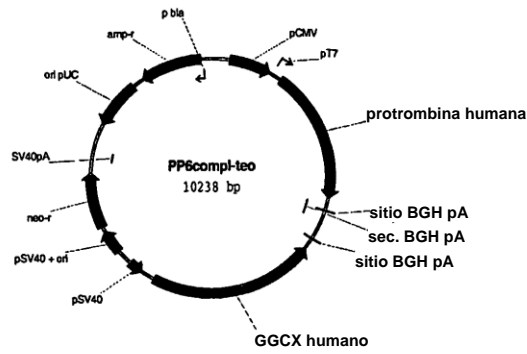


Figura 3a

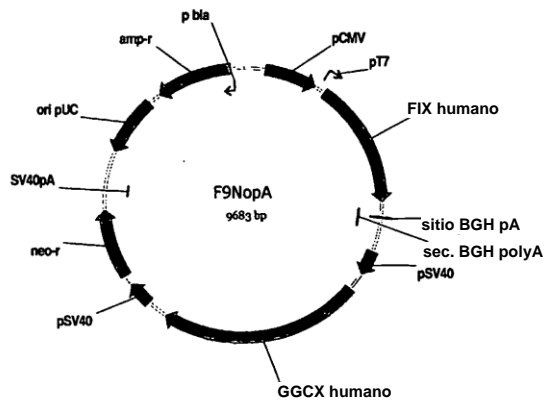


Figura 3b

