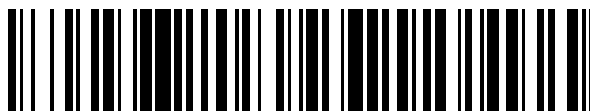


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 573 456**

51 Int. Cl.:

C07H 21/00	(2006.01)
C07H 21/02	(2006.01)
C07H 21/04	(2006.01)
C12N 15/63	(2006.01)
A61K 48/00	(2006.01)
A61K 39/39	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.02.2007 E 07752109 (4)**

97 Fecha y número de publicación de la concesión europea: **06.04.2016 EP 1996238**

54 Título: **Vectores adenovirales quiméricos y ARNbc como agonista de TLR3**

30 Prioridad:

28.02.2006 US 778026 P
 19.05.2006 US 801645 P
 22.05.2006 US 802992 P
 04.08.2006 US 821492 P
 22.09.2006 US 846658 P
 28.09.2006 US 848195 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
08.06.2016

73 Titular/es:

VAXART, INC. (100.0%)
600 Townsend Street Suite 120E
San Francisco, CA 94103, US

72 Inventor/es:

TUCKER, SEAN N.

74 Agente/Representante:

PONS ARIÑO, Ángel

Observaciones :

Véase nota informativa (Remarks) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 573 456 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

Vectores adenovirales quiméricos y ARNbc como agonista de TLR3

5 Antecedentes de la invención

Las vacunas son medios importantes para prevenir y/o tratar varias enfermedades y trastornos (por ejemplo, infección viral, infección bacteriana y cáncer). Las vacunas basadas en ácidos nucleicos tienen varias ventajas con respecto a las vacunas de proteína o vivas atenuadas. La introducción de un ácido nucleico que expresa un antígeno en una célula diana permite el rápido desarrollo de vacuna que genera una respuesta inmunitaria contra un antígeno de interés. Para las vacunas de proteína necesita desarrollarse un método de purificación de proteínas eficaz y eficiente cada vez que se crea una nueva vacuna. Para las vacunas vivas necesita identificarse un método de atenuación que no detenga completamente el crecimiento del patógeno, que ya ha demostrado ser completamente seguros en seres humanos. El desarrollo de metodologías de purificación y de atenuación de proteínas son procesos que requieren extremadamente mucho tiempo. A diferencia, la mayoría de las vacunas basadas en ácidos nucleicos pueden fabricarse muy rápidamente usando las mismas técnicas de fabricación cada vez con solo un cambio rápido en el ácido nucleico que codifica el antígeno de interés. El adenovirus incompetente en la replicación es un sistema de vacuna basado en ácidos nucleicos que se prepara rápidamente, predeciblemente y económicamente a alto título [Polo, J. M. y Dubensky, T. W., Jr., Drug Discov Today, 7(13), 719-727 (2002)]. Sin embargo, la eficiencia de la respuesta específica de antígeno tras la administración de vectores adenovirales conocidos en la técnica es baja. El documento US6511845 describe inmunización por administración intranasal de un vector de adenovirus que expresa un antígeno, seguido de una o más administraciones de refuerzo intranasales o intramusculares de dicho adenovirus recombinante. El documento WO2005014038 describe un efecto adyuvante de ARNbc sobre la inmunidad mucosa inducida contra un antígeno de subunidad o un antígeno inactivado de un patógeno. Así, hay una necesidad en la materia de nuevos vectores adenovirales que puedan usarse para provocar eficientemente una respuesta inmunitaria contra un antígeno de interés. La presente invención cumple estas y otras necesidades.

Sumario de la invención

La invención en su sentido más amplio es como se define en las reivindicaciones independientes.

Una realización de la invención proporciona un vector de expresión adenoviral quimérico, comprendiendo dicho vector un casete de expresión que comprende los siguientes elementos: (a) un primer promotor unido operativamente a un ácido nucleico que codifica un agonista de receptor tipo Toll (TLR)-3, en el que el agonista de TLR-3 es ARNbc heterólogo y en el que el ácido nucleico que codifica el agonista de TLR-3 comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NOS: 3, 4, 5, 8, 9, 10, 11 y 12; y (b) un segundo promotor unido operativamente a un ácido nucleico que codifica un polipéptido heterólogo inmunogénico. En algunas realizaciones, el polipéptido heterólogo está seleccionado de un polipéptido de la envuelta del VIH (por ejemplo, gp41, gp120 o gp160) y polipéptido HA de la gripe. En algunas realizaciones, el primer y segundo promotores son los mismos. En algunas realizaciones, el primer y segundo promotores son diferentes. En algunas realizaciones, los promotores están seleccionados del promotor de beta-actina y un promotor del CMV. La invención también proporciona composiciones inmunogénicas que comprenden el vector de expresión.

Otra realización de la invención proporciona un vector de expresión adenoviral quimérico, comprendiendo dicho vector un casete de expresión que comprende los siguientes elementos:

- (a) un primer promotor unido operativamente a un ácido nucleico que codifica un agonista de receptor 3 tipo Toll (TLR-3), en el que el agonista de TLR-3 es ARNbc heterólogo; y
- (b) un segundo promotor unido operativamente a un ácido nucleico que codifica un polipéptido heterólogo inmunogénico para su uso en un método de tratamiento que provoca una respuesta inmunitaria, comprendiendo dicho método administrar una cantidad inmunogénicamente eficaz del vector a un sujeto mamífero tal como un ser humano, en el que la respuesta inmunitaria se dirige contra el polipéptido heterólogo, y en el que la vía de administración está seleccionada del grupo que consiste en oral, intranasal, y mucosa tal como vaginal. El sujeto mamífero es, por ejemplo, un roedor tal como un ratón, una rata, o una cobaya o un primate tal como un chimpancé, un macaco rhesus, o un ser humano. En algunas realizaciones, el polipéptido heterólogo se expresa en una célula seleccionada de una célula dendrítica, una célula de micropliegue y una célula epitelial intestinal.

Otra realización de la invención proporciona una composición inmunogénica, comprendiendo dicha composición:

- (a) un vector de expresión adenoviral quimérico que comprende un promotor unido operativamente a un ácido nucleico que codifica un polipéptido heterólogo inmunogénico;
- (b) un agonista de TLR-3, en el que el agonista de TLR-3 es ARNbc heterólogo; y
- (c) un vehículo farmacéuticamente aceptable;

para su uso en un método de tratamiento que provoca una respuesta inmunitaria, comprendiendo dicho método

5 administrar una cantidad inmunogénicamente eficaz de la composición a un sujeto mamífero tal como un ser humano, en el que la respuesta inmunitaria se dirige contra el polipéptido heterólogo, y en el que la vía de administración está seleccionada del grupo que consiste en oral, intranasal, y mucosa tal como vaginal. El sujeto mamífero es, por ejemplo, un roedor tal como un ratón, una rata, o una cobaya o un primate tal como un chimpancé, un macaco rhesus, o un ser humano.

La divulgación también proporciona un ácido nucleico aislado que comprende la secuencia expuesta en SEQ ID NOS: 1, 2, 6, 7, 13, 14, 15, 16 o 17.

10 **Breve descripción de los dibujos**

15 La Figura 1 ilustra datos que demuestran que un vector adenoviral quimérico de la invención (es decir, DS1) en combinación con un agonista de TLR-3 es más eficaz que un vector adenoviral estándar (es decir, rAd5) en inducir una respuesta inmunitaria específica de antígeno tras la administración del vector oral. La Figura 1A ilustra datos que representan el título de anticuerpos para la proteína de la envuelta del VIH (es decir, gp120) 3 semanas tras la administración oral de los vectores adenovirales. La Figura 1B ilustra datos que representan el título de anticuerpos para la proteína de la envuelta del VIH (es decir, gp120) 6 semanas tras la administración oral de los vectores adenovirales.

20 La Figura 2 ilustra datos que demuestran que un vector adenoviral quimérico de la invención (es decir, DS1b o DS1c) en combinación con un agonista de TLR-3 es más eficaz en inducir una respuesta inmunitaria específica de antígeno que un vector adenoviral estándar (es decir, rAd5). La Figura 2A ilustra datos que representan el título de IgG anti-GFP 3 semanas tras la administración por vía oral de los vectores. La Figura 2B ilustra datos que representan la respuesta de linfocitos T CD8+ a GFP 10 semanas tras la administración del vector a 0, 4 y 8 semanas. La Figura 2C ilustra datos que representan el título de anticuerpos anti-HA 3 semanas tras la administración por vía oral de los vectores.

25 La Figura 3 ilustra datos que demuestran que los vectores adenovirales quiméricos de la invención son superiores para provocar respuestas inmunitarias cuando se administran no parenteralmente.

30 La Figura 3A ilustra datos que representan el título de anticuerpos anti-gp120 3 semanas tras la administración intramuscular de DS1. La Figura 3B ilustra datos que representan el título de anticuerpos anti-HA tres semanas tras la administración intranasal de DS1c.

La Figura 4 ilustra datos que demuestran que los agonistas de ligandos TLR-3 expresados pueden inducir la activación de células presentadoras de antígeno. La Figura 4A ilustra datos que representan activación de células dendríticas por el agonista de TLR-3 de ARNbc expresado luc1. La Figura 4B ilustra datos que representan la activación de células dendríticas por los agonistas de TLR-3 de ARNbc expresados luc1 y ml.

35 La Figura 5 es una ilustración gráfica de los vectores adenovirales quiméricos de la invención, es decir, vectores adenovirales quiméricos que comprenden ácidos nucleicos que codifican agonistas de TLR-3 de ARNbc expresados.

40 La Figura 6 ilustra datos que demuestran que los vectores adenovirales quiméricos de la invención son eficaces en inducir una respuesta inmunitaria específica de antígeno tras la administración oral. La Figura 6 ilustra datos que representan el título de anticuerpos anti-gp120 3 semanas tras la administración por vía oral de un adenoviral quimérico que comprende una secuencia de ácidos nucleicos que codifica el agonista de TLR-3 de ARNbc luc1.

45 La Figura 7 ilustra datos que demuestran que agonistas de TLR-7/8 tienen mala eficacia en inducir una respuesta inmunitaria específica de antígeno.

La Figura 8 ilustra datos que demuestran que vectores adenovirales quiméricos de la invención son eficaces en inducir una respuesta inmunitaria específica de antígeno tras la administración oral. La Figura 8A ilustra datos que representan el título de anticuerpos anti-HA 4 semanas tras la administración por vía oral de un adenoviral quimérico que comprende una secuencia de ácidos nucleicos que codifica el agonista de TLR-3 de ARNbc luc1.

50 La Figura 8B ilustra datos que representan el título de anticuerpos anti-HA 4 semanas o 7 semanas tras la administración de un adenoviral quimérico que comprende una secuencia de ácidos nucleicos que codifica el agonista de TLR-3 de ARNbc luc1. La Figura 8C ilustra datos que representan el título de anticuerpos anti-HA 3 semanas tras la administración oral o intranasal de un adenoviral quimérico que comprende una secuencia de ácidos nucleicos que codifica el agonista de TLR-3 de ARNbc luc1.

55 **Breve descripción de las secuencias**

SEQ ID NO: 1 expone la secuencia de nucleótidos para el vector adenoviral quimérico DS1.

SEQ ID NO: 2 expone la secuencia de nucleótidos para el vector adenoviral quimérico DS2.

60 SEQ ID NO: 3 expone una secuencia de nucleótidos que codifica un agonista de TLR-3.

SEQ ID NO: 4 expone una secuencia de nucleótidos que codifica un agonista de TLR-3.

SEQ ID NO: 5 expone una secuencia de nucleótidos que codifica un agonista de TLR-3.

SEQ ID NO: 6 expone una secuencia de nucleótidos para un vector adenoviral quimérico que comprende un ácido nucleico que codifica HA de la gripe y un ácido nucleico que codifica un agonista de TLR-3 (luc), en la que la HA de la gripe y el agonista de TLR-3 están en la misma orientación.

65 SEQ ID NO: 7 se expone una secuencia de nucleótidos para un vector adenoviral quimérico que comprende un ácido nucleico que codifica HA de la gripe y un ácido nucleico que codifica un agonista de TLR-3 (luc), en la que

la HA de la gripe y el agonista de TLR-3 están en la orientación opuesta.

SEQ ID NO: 8 se expone una secuencia de nucleótidos que codifica un agonista de TLR-3 de ARN de horquilla corta. Porciones complementarias de la secuencia se muestran en mayúsculas y la secuencia de conector se muestra en minúsculas.

SEQ ID NO: 9 se expone una secuencia de nucleótidos que codifica un agonista de TLR-3 de ARN de horquilla corta (g1). Porciones complementarias de la secuencia se muestran en mayúsculas y la secuencia de conector se muestra en minúsculas.

SEQ ID NO: 10 se expone una secuencia de nucleótidos que codifica un agonista de TLR-3 de ARN de horquilla corta (luc). Porciones complementarias de la secuencia se muestran en mayúsculas y la secuencia de conector se muestra en minúsculas.

SEQ ID NO: 11 se expone una secuencia de nucleótidos que codifica un agonista de TLR-3 de ARN de horquilla corta (ml). Porciones complementarias de la secuencia se muestran en mayúsculas y la secuencia de conector se muestra en minúsculas.

SEQ ID NO: 12 se expone una secuencia de nucleótidos que codifica un agonista de TLR-3 de ARN de horquilla corta. Porciones complementarias de la secuencia se muestran en mayúsculas y la secuencia de conector se muestra en minúsculas.

SEQ ID NO: 13 se expone la secuencia de nucleótidos para el vector adenoviral quimérico DS 1c. La secuencia comprende un nucleótido que codifica HA (PR8/34).

SEQ ID NO: 14 se expone la secuencia de nucleótidos para el vector adenoviral quimérico DS2beta-luc. El vector comprende una secuencia que codifica el agonista de TLR-3 luc bajo el control del promotor de beta-actina. El vector también comprende sitios de clonación abiertos para la inserción de secuencia(s) de ácidos nucleicos que codifica(n) un antígeno de interés.

SEQ ID NO: 15 se expone la secuencia de nucleótidos para el vector adenoviral quimérico DS2C-luc. El vector comprende una secuencia que codifica el agonista de TLR-3 luc bajo el control de un promotor del CMV. El vector también comprende sitios de clonación abiertos para la inserción de secuencia(s) de ácidos nucleicos que codifica(n) un antígeno de interés.

SEQ ID NO: 16 se expone la secuencia de nucleótidos para el vector pShuttle que comprende una secuencia de ácidos nucleicos que codifica el agonista de TLR-3 luc bajo el control de un promotor del CMV y una secuencia de ácidos nucleicos que codifica HA (gripe aviar) bajo el control de promotor del CMV separado.

SEQ ID NO: 17 se expone la secuencia de nucleótidos para el vector adenoviral quimérico ND1.1 214. El ácido nucleico que codifica el antígeno heterólogo está en negrita y está flanqueado por un sitio de reconocimiento Cla I sobre el extremo 5' y un sitio de reconocimiento Not I sobre el extremo 3'. La secuencia de ácidos nucleicos que codifica los agonistas de TLR-3 está en cursiva, con la secuencia de conector en negrita.

Descripción detallada de la invención

La invención en su sentido más amplio es como se define en las reivindicaciones independientes.

I. Introducción

La presente invención proporciona vectores adenovirales quiméricos novedosos que pueden administrarse no parenteralmente para provocar una respuesta inmunitaria contra un antígeno de interés. Los vectores adenovirales quiméricos de la invención comprenden un ácido nucleico que codifica un polipéptido heterólogo y un ácido nucleico que codifica un agonista de TLR-3. Los vectores adenovirales quiméricos provocan respuestas inmunitarias fuertes y eficaces específicas para el polipéptido heterólogo, particularmente cuando se administran por una vía no parenteral (por ejemplo, por vía oral, intranasalmente, o mucosamente).

La invención se basa en el sorprendente descubrimiento de que la administración de agonistas de TLR-3 de ARNbc son adyuvantes eficaces cuando se administran conjuntamente con vectores virales. En realidad, el uso de ARNbc como adyuvante para vectores virales sería contradictorio considerando que la principal utilidad propuesta del mimético de ARNbc poli I:C era como agente antiviral [Nemes, et al., Proc Soc Exp Biol Med. (1969) 132:776; Schafer, et al, Nature. (1970) 226:449; Fenje, et al, Nature (1970) 226:171.].

II. Definiciones

El término "quimérico" o "recombinante", como se usa en el presente documento con referencia, por ejemplo, a un ácido nucleico, proteína o vector, indica que el ácido nucleico, proteína o vector se ha modificado por la introducción de un ácido nucleico heterólogo o proteína o la alteración de un ácido nucleico nativo o proteína. Así, por ejemplo, vectores quiméricos y recombinantes incluyen secuencias de ácidos nucleicos que no se encuentran dentro de la forma nativa (no quimérica o no recombinante) del vector. Un vector de expresión adenoviral quimérico se refiere a un vector de expresión adenoviral que comprende una secuencia de ácidos nucleicos que codifica un polipéptido heterólogo.

Un "vector de expresión" es una construcción de ácidos nucleicos, generada recombinantemente o sintéticamente, con una serie de elementos de ácido nucleico especificados que permiten la transcripción de un ácido nucleico particular en una célula huésped. El vector de expresión puede ser parte de un plásmido, virus o fragmento de ácido

nucleico. Normalmente, el vector de expresión incluye un ácido nucleico que va a transcribirse unido operativamente a un promotor.

Los términos “promotor” y “secuencia de control de la expresión” se usan en el presente documento para referirse a una matriz de secuencias de control de ácido nucleico que dirigen la transcripción de un ácido nucleico. Como se usa en el presente documento, un promotor incluye secuencias de ácidos nucleicos necesarias cerca del sitio de inicio de la transcripción, tales como en el caso de un promotor tipo polimerasa II, un elemento TATA. Un promotor también opcionalmente incluye elementos potenciadores o represores distales, que pueden localizarse nada menos que varios miles de pares de bases del sitio de inicio de la transcripción. Los promotores incluyen promotores constitutivos e inducibles. Un promotor “constitutivo” es un promotor que es activo bajo la mayoría de las condiciones ambientales y de desarrollo. Un promotor “inducible” es un promotor que es activo bajo regulación ambiental o de desarrollo. El término “operativamente ligado” se refiere a un enlace funcional entre una secuencia de control de la expresión de ácido nucleico (tal como un promotor, o matriz de sitios de unión al factor de transcripción) y una segunda secuencia de ácidos nucleicos, en la que la secuencia de control de la expresión dirige la transcripción del ácido nucleico correspondiente a la segunda secuencia.

Los términos “agonista de TLR” o “agonista de receptor tipo Toll”, como se usa en el presente documento, se refiere a un compuesto que se une y estimula un receptor tipo Toll que incluye, por ejemplo, TLR-2, TLR-3, TLR-6, TLR-7 o TLR-8. Los agonistas de TLR se revisan en MacKichan, IAVI Report. 9:1-5 (2005) y Abreu et al., J Immunol, 174(8), 4453-4460 (2005). Los agonistas inducen la transducción de señales tras la unión a su receptor.

Los términos “agonista de TLR-3” o “agonista de receptor 3 tipo Toll”, como se usan en el presente documento, se refieren a un compuesto que se une y estimula TLR-3. Se ha identificado que los agonistas de TLR-3 incluye ARN bicatenario, ARNbc viralmente derivado, varios análogos químicamente sintetizados para ARN bicatenario que incluyen poli-inosina-ácido policitidílico (poli I:C) - ácido poliadenílico-poliuridílico (poli A:U) y poli I:poli C, y anticuerpos (o reticulación de anticuerpos) para TLR-3 que conducen a la producción de IFN-beta [Matsumoto, M, et al, Biochem Biophys Res Commun 24:1364 (2002), de Bouteiller, et al, J Biol Chem 18:38133-45 (2005)]. Los agonistas de TLR-3 también incluyen ARNbc expresados (por ejemplo, ARNbc codificado por un ácido nucleico que comprende una secuencia expuesta en SEQ ID NOS: 3, 7, 8, 9, 10, 11 o 12).

Los términos “agonista de TLR-7/8” o “agonista de receptor 7/8 tipo Toll”, como se usan en el presente documento, se refieren a un compuesto que se une y estimula tanto los receptores de TLR-7 como de TLR-8; estos receptores reconocen varios de los mismos ligandos. Se han identificado varios agonistas de TLR-7/8 tales como ARN monocatenario viral, imiquimod, loxoribina, ácido poliuridílico, o resiquimod.

El término “heterólogo”, cuando se usa con referencia a porciones de un ácido nucleico, indica que el ácido nucleico comprende dos o más subsecuencias que no se encuentran en la misma relación entre sí en la naturaleza. Por ejemplo, el ácido nucleico normalmente se produce recombinantemente, teniendo dos o más secuencias de genes sin relacionar dispuestos para hacer un nuevo ácido nucleico funcional, por ejemplo, un promotor de una fuente y una región codificante de otra fuente. Similarmente, una proteína heteróloga indica que la proteína comprende dos o más subsecuencias que no se encuentran en la misma relación entre sí en la naturaleza (por ejemplo, una proteína de fusión).

Los términos “ácido nucleico” y “polinucleótido” se usan indistintamente en el presente documento para referirse a desoxirribonucleótidos o ribonucleótidos y polímeros de los mismos en tanto forma mono como bicatenaria. El término engloba ácidos nucleicos que contienen análogos de nucleótidos conocidos o residuos o enlaces de esqueleto modificados, que son sintéticos, que existen de forma natural, y que no existen de forma natural, que tienen propiedades de unión similares como el ácido nucleico de referencia, y que son metabolizados de un modo similar a los nucleótidos de referencia. Ejemplos de tales análogos incluyen, sin limitación, fosforotioatos, fosforamidatos, fosfonatos de metilo, fosfonatos de metilo quirales, 2-O-metilribonucleótidos, ácidos nucleicos peptídicos (PNA).

A menos que se indique lo contrario, una secuencia de ácidos nucleicos particular también engloba variantes conservativamente modificadas de la misma (por ejemplo, sustituciones de codón degenerado) y secuencias complementarias, además de la secuencia explícitamente indicada. Específicamente, las sustituciones de codón degenerado pueden lograrse generando secuencias en las que la tercera posición de uno o más codones seleccionados (o todas) está sustituida con residuos de base mixta y/o desoxiinosina (Batzer et al., Nucleic Acid Res. 19:5081 (1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)). El término ácido nucleico se usa indistintamente con gen, ADNc, ARNm, oligonucleótido y polinucleótido.

Antígeno se refiere a una proteína o parte de una cadena de polipéptidos que puede ser reconocida por receptores de linfocito T y/o anticuerpos. Normalmente, los antígenos se derivan de proteínas bacterianas, virales o fúngicas.

Una “dosis o cantidad inmunogénicamente eficaz” de las composiciones de la presente invención es una cantidad que provoca o modula una respuesta inmunitaria específica para el polipéptido heterólogo. Respuestas inmunitarias incluyen respuestas inmunitarias humorales y respuestas inmunitarias celulares. Una composición inmunogénica

puede usarse terapéuticamente o profilácticamente para tratar o prevenir enfermedad en cualquier etapa.

Las “respuestas inmunitarias humorales” están mediadas por componentes libres de células de la sangre, es decir, plasma o suero; la transferencia de suero o plasma de un individuo a otro transfiere inmunidad.

Las “respuestas inmunitarias celulares” están mediadas por linfocitos específicos de antígeno; la transferencia de los linfocitos específicos de antígeno de un individuo a otro transfiere inmunidad.

Una “dosis terapéutica” o “cantidad terapéuticamente eficaz” o “cantidad eficaz” de un vector adenoviral quimérico o una composición que comprende un vector adenoviral quimérico es una cantidad de vector o composición que comprende el vector que previene, alivia, calma o reduce la gravedad de síntomas de enfermedades y trastornos asociados a la fuente del polipéptido heterólogo (por ejemplo, un virus, bacteria, un parásito, o un cáncer).

Anticuerpo se refiere a un polipéptido codificado por un gen de inmunoglobulina o fragmentos del mismo, que se une específicamente a y reconoce un antígeno. Los genes de inmunoglobulina reconocidos incluyen los genes de las regiones constantes kappa, lambda, alfa, gamma, delta, épsilon y mu, además de infinidad de genes de la región variable de inmunoglobulina. Las cadenas ligeras se clasifican como tanto kappa como lambda. Las cadenas pesadas se clasifican como gamma, mu, alfa, delta o épsilon, que a su vez definen las clases de inmunoglobulina, IgG, IgM, IgA, IgD e IgE, respectivamente.

Linfocitos T se refiere a una clase particular de linfocitos que expresan un receptor específico (receptor de linfocito T) codificado por una familia de genes. Los genes del receptor de linfocito T reconocidos incluyen loci alfa, beta, delta, y gamma, y los receptores de linfocito T normalmente (pero no universalmente) reconocen una combinación de MHC más un péptido corto.

Respuesta inmunitaria adaptativa se refiere al reconocimiento de linfocito T y/o anticuerpo de antígeno.

Células presentadoras de antígeno (APC), como se usa en el presente documento, se refiere a células que son capaces de presentar péptidos inmunogénicos o fragmentos de los mismos a linfocitos T para activar o potenciar una respuesta inmunitaria. Las APC incluyen células dendríticas, macrófagos, linfocitos B, monocitos y otras células que pueden manipularse para ser APC eficientes. Tales células pueden, pero no necesitan, ser genéticamente modificadas para aumentar la capacidad para presentar el antígeno, para mejorar la activación y/o mantenimiento de la respuesta de linfocitos T, para tener efectos antitumorales por sí mismas y/o para ser inmunológicamente compatibles con el receptor (es decir, mismo haplotipo de HLA). Las APC pueden aislarse de cualquiera de una variedad de fluidos biológicos y órganos que incluyen médula ósea, sangre periférica, tumor y tejidos peritumorales, y pueden ser células autólogas, alógenas, singénicas o xenógenas. Las APC normalmente utilizan un receptor del sitio de histocompatibilidad mayor (MHC) para presentar polipéptidos cortos a linfocitos T.

Adyuvante es un potenciador de la respuesta inmunitaria no específico. Adyuvantes adecuados incluyen, por ejemplo, toxina del cólera, monofosforil lipido A (MPL), adyuvante completo de Freund, adyuvante incompleto de Freund, Quil A y Al(OH). Los adyuvantes también pueden ser aquellas sustancias que producen activación de APC y presentación potenciada de linfocitos T mediante moléculas de señalización secundaria como receptores tipo Toll. Ejemplos de receptores tipo Toll incluyen los receptores que reconocen ARN bicatenario, flagelos bacterianos, LPS, ADN de CpG y lipopéptido bacteriano (revisado recientemente en [Abreu et al, J Immunol, 174(8), 4453-4460 (2005)]).

Los términos “polipéptido”, “péptido” y “proteína” se usan indistintamente en el presente documento para referirse a un polímero de residuos de aminoácidos. Los términos se aplican a polímeros de aminoácidos en los que uno o más residuos de aminoácido es un mimético químico artificial de un aminoácido correspondiente que existe de forma natural, además de a polímeros de aminoácidos que existen de forma natural y polímero de aminoácido que no existe de forma natural.

El término “aminoácido” se refiere a aminoácidos que existen de forma natural y sintéticos, además de a análogos de aminoácidos y miméticos de aminoácido que funcionan de un modo similar a los aminoácidos que existen de forma natural. Los aminoácidos que existen de forma natural son aquellos codificados por el código genético, además de aquellos aminoácidos que son después modificados, por ejemplo, hidroxiprolina, γ-carboxiglutamato y O-fosfoserina. Análogos de aminoácido se refiere a compuestos que tienen la misma estructura química básica que un aminoácido que existe de forma natural, es decir, un carbono γ que está unido a un hidrógeno, un grupo carboxilo, un grupo amino y un grupo R, por ejemplo, homoserina, norleucina, sulfóxido de metionina, metionina metil sulfonio. Tales análogos tienen grupos R modificados (por ejemplo, norleucina) o esqueletos de péptido modificados, pero retienen la misma estructura química básica que un aminoácido que existe de forma natural. Miméticos de aminoácido se refiere a compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero que funciona de un modo similar a un aminoácido que existe de forma natural.

Los aminoácidos pueden denominarse en el presente documento por tanto sus símbolos de tres letras comúnmente conocidos como por los símbolos de una letra recomendados por la IUPAC-IUB Biochemical Nomenclature

Commission. Los nucleótidos, asimismo, pueden denominarse por sus códigos de una letra comúnmente aceptados.

“Variantes conservativamente modificadas” se aplica a tanto secuencias de aminoácidos como de ácidos nucleicos. Con respecto a secuencias de ácidos nucleicos particulares, variantes conservativamente modificadas se refiere a aquellos ácidos nucleicos que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o en las que el ácido nucleico no codifica una secuencia de aminoácidos, para secuencias esencialmente idénticas. Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier proteína dada. Por ejemplo, todos los codones GCA, GCC, GCG y GCU codifican el aminoácido alanina. Así, en cualquier posición en la que una alanina esté especificada por un codón, el codón puede alterarse a cualquiera de los codones correspondientes descritos sin alterar el polipéptido codificado. Tales variaciones de ácido nucleico son “variaciones silenciosas”, que son una especie de variaciones conservativamente modificadas. Cada secuencia de ácidos nucleicos en el presente documento que codifica un polipéptido también describe cada posible variación silenciosa del ácido nucleico. Un experto reconocerá que cada codón en un ácido nucleico (excepto AUG, que es generalmente el único codón para metionina, y TGG, que es generalmente el único codón para triptófano) puede modificarse para dar una molécula funcionalmente idéntica. Por consiguiente, cada variación silenciosa de un ácido nucleico que codifica un polipéptido está implícita en cada secuencia descrita.

En cuanto a las secuencias de aminoácidos, un experto reconocerá que sustituciones, deleciones o adiciones individuales a una secuencia de ácido nucleico, péptido, polipéptido o proteína que altera, añade o delecciona un único aminoácido o un pequeño porcentaje de aminoácidos en la secuencia codificada es una “variante conservativamente modificada” en la que la alteración produce la sustitución de un aminoácido con un aminoácido químicamente similar. Tablas de sustituciones conservativas que proporcionan aminoácidos funcionalmente similares son muy conocidas en la técnica. Tales variantes conservativamente modificadas son además de y no excluyen variantes polimórficas, homólogos entre especies y alelos de la invención.

Los ocho siguientes grupos contienen cada uno aminoácidos que son sustituciones conservativas para los otros:

- 1) Alanina (A), Glicina (G);
- 2) Ácido aspártico (D), Ácido glutámico (E);
- 3) Asparagina (N), Glutamina (Q);
- 4) Arginina I, Lisina (K);
- 5) Isoleucina (I), Leucina (L), Metionina (M), Valina (V);
- 6) Fenilalanina (F), Tirosina (Y), Triptófano (W);
- 7) Serina (S), Treonina (T); y
- 8) Cisteína (C), Metionina (M)

(véase, por ejemplo, Creighton, Proteins (1984)).

La expresión “se hibrida selectivamente (o específicamente) con” se refiere a la unión, duplexado o hibridación de una molécula sola a una secuencia de nucleótidos particular bajo condiciones de hibridación rigurosas cuando esa secuencia está presente en una mezcla compleja (por ejemplo, ADN o ARN celular total o de biblioteca).

La expresión “condiciones de hibridación rigurosas” se refiere a condiciones bajo las que una sonda se hibridará con su subsecuencia diana, normalmente en una mezcla compleja de ácido nucleico, pero no con otras secuencias. Condiciones rigurosas son dependientes de secuencia y serán diferentes en diferentes circunstancias. Las secuencias más largas se hibridan específicamente a mayores temperaturas. Una amplia guía a la hibridación de ácidos nucleicos se encuentra en Tijssen, Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes, “Overview of principles of hybridization and the strategy of nucleic acid assays” (1993). Generalmente, las condiciones rigurosas están seleccionadas para ser aproximadamente 5-10 °C inferiores al punto de fusión térmico T_m para la secuencia específica a una fuerza iónica definida Ph . La T_m es la temperatura (bajo fuerza iónica definida, Ph , y concentración nucleica) a la que el 50 % de las sondas complementarias a la diana se hibridan con la secuencia diana en equilibrio (como las secuencias diana están presentes en exceso, a T_m , el 50 % de las sondas están ocupadas en equilibrio). Condiciones rigurosas serán aquellas en las que la concentración de sales es inferior a ión sodio aproximadamente 1,0 M, normalmente concentración de ión sodio aproximadamente 0,01 a 1,0 M (u otras sales) a pH 7,0 a 8,3 y la temperatura es al menos aproximadamente 30 °C para sondas cortas (por ejemplo, 10 a 50 nucleótidos) y al menos aproximadamente 60 °C para sondas largas (por ejemplo, superior a 50 nucleótidos). También pueden lograrse condiciones rigurosas con la adición de agentes desestabilizantes tales como formamida. Para la hibridación selectiva o específica, una señal positiva es al menos dos veces el fondo, opcionalmente 10 veces la hibridación de fondo. Condiciones de hibridación rigurosas a modo de ejemplo pueden ser las siguientes: 50 % de formamida, 5x SSC, y 1 % de SDS, incubar a 42 °C, o, 5x SSC, 1 % de SDS, incubar a 65 °C, con lavado en 0,2x SSC, y 0,1 % de SDS a 65 °C.

Ácidos nucleicos que no se hibridan entre sí bajo condiciones rigurosas son todavía sustancialmente idénticos si los polipéptidos que codifican son sustancialmente idénticos. Esto se produce, por ejemplo, cuando una copia de un ácido nucleico se crea usando la máxima degeneración de codones permitida por el código genético. En tales casos, los ácidos nucleicos normalmente se hibridan bajo condiciones de hibridación moderadamente rigurosas. A modo de

ejemplo “condiciones de hibridación moderadamente rigurosas” incluyen una hibridación en un tampón de 40 % de formamida, NaCl 1 M, 1 % de SDS a 37 °C, y un lavado en 1X SSC a 45 °C. Una hibridación positiva es al menos dos veces el fondo. Aquellos expertos habituales reconocerán fácilmente que pueden utilizarse condiciones alternativas de hibridación y de lavado para proporcionar condiciones de rigurosidad similar.

“Anticuerpo” se refiere a un polipéptido que comprende una región estructural de un gen de inmunoglobulina o fragmentos del mismo que se une específicamente a y reconoce un antígeno. Los genes de inmunoglobulina reconocidos incluyen los genes de las regiones constantes kappa, lambda, alfa, gamma, delta, épsilon y mu, además de infinidad de genes de la región variable de inmunoglobulina. Las cadenas ligeras se clasifican como tanto kappa como lambda. Las cadenas pesadas se clasifican como gamma, mu, alfa, delta o épsilon, que a su vez definen las clases de inmunoglobulina, IgG, IgM, IgA, IgD e IgE, respectivamente.

Una unidad estructural de inmunoglobulina (anticuerpo) a modo de ejemplo comprende un tetrámero. Cada tetrámero está compuesto por dos pares idénticos de cadenas de polipéptidos, teniendo cada par una cadena “ligera” (aproximadamente 25 kDa) y una “pesada” (aproximadamente 50-70 kDa). El extremo N de cada cadena define una región variable de aproximadamente 100 a 110 o más aminoácidos principalmente responsables del reconocimiento del antígeno. Los términos cadena ligera variable (V_L) y cadena pesada variable (V_H) se refieren a estas cadenas ligeras y pesadas, respectivamente.

La expresión “se une específicamente (o selectivamente)” a un anticuerpo o “específicamente (o selectivamente) inmunorreactivo con”, cuando se refiere a una proteína o péptido, se refiere a una reacción de unión que es determinante de la presencia de la proteína en una población heterogénea de proteínas y otros biológicos. Así, bajo condiciones de inmunoensayo designadas, los anticuerpos especificados se unen a una proteína particular al menos dos veces el fondo y no se unen sustancialmente en una cantidad significativa a otras proteínas presentes en la muestra. La unión específica a un anticuerpo bajo tales condiciones puede requerir un anticuerpo que se selecciona por su especificidad por una proteína particular. Por ejemplo, pueden seleccionarse anticuerpos policlonales producidos contra proteínas de fusión para obtener solo aquellos anticuerpos policlonales que son específicamente inmunorreactivos con proteína de fusión y no con componentes individuales de las proteínas de fusión. Esta selección puede lograrse restando anticuerpos que reaccionan de forma cruzada con los antígenos individuales. Puede usarse una variedad de formatos de inmunoensayo para seleccionar anticuerpos específicamente inmunorreactivos con una proteína particular. Por ejemplo, se usan rutinariamente inmunoensayos de ELISA en fase sólida para seleccionar anticuerpos específicamente inmunorreactivos con una proteína (véase, por ejemplo, Harlow & Carril, *Antibodies, A Laboratory Manual* (1988), para una descripción de formatos y condiciones de inmunoensayo que pueden usarse para determinar inmunorreactividad específica). Normalmente, una reacción específica o selectiva será al menos dos veces la señal de fondo o ruido y más normalmente superior a 10 a 100 veces el fondo.

Los polinucleótidos pueden comprender una secuencia nativa (es decir, una secuencia endógena que codifica un polipéptido individual o ARNbc o una porción del mismo) o pueden comprender una variante de una secuencia tal. Las variantes de polinucleótidos pueden contener una o más sustituciones, adiciones, deleciones y/o inserciones de forma que la actividad biológica del polipéptido codificado no se reduzca, con respecto a un polipéptido que comprende antígenos nativos. Las variantes de polinucleótidos pueden contener una o más sustituciones, adiciones, deleciones y/o inserciones de forma que no se reduzca la actividad de agonista de TLR-3 del ARNbc codificado, con respecto a un ARNbc que no contiene las sustituciones, adiciones, deleciones y/o inserciones. Las variantes presentan preferentemente al menos aproximadamente el 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, con una secuencia de polinucleótidos que codifica un polipéptido nativo o una porción del mismo o con una secuencia de polinucleótidos que codifica un ARNbc con actividad de agonista de TLR-3.

Los términos “idénticos” o porcentaje de “identidad”, en el contexto de dos o más ácidos nucleicos (por ejemplo, un ARNbc que es un agonista de TLR-3) o secuencias de polipéptidos, se refieren a dos o más secuencias o subsecuencias que son las mismas o tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son los mismos (es decir, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o más de identidad con respecto a una región especificada), cuando se comparan y alinean para la correspondencia máxima sobre una ventana de comparación, o región designada como se mide usando uno de los siguientes algoritmos de comparación de secuencias o por alineamiento manual e inspección visual. Se dice entonces que tales secuencias son “sustancialmente idénticas”. Esta definición también se refiere al complemento de una secuencia de prueba. Opcionalmente, la identidad existe sobre una región que tiene al menos aproximadamente 10 a aproximadamente 100, aproximadamente 20 a aproximadamente 75, aproximadamente 30 a aproximadamente 50 aminoácidos o nucleótidos de longitud.

Para comparación de secuencias, normalmente una secuencia actúa de una secuencia de referencia, con la que se comparan las secuencias de prueba. Si se usa un algoritmo de comparación de secuencias, las secuencias de prueba y de referencia se entran en un ordenador, se designan coordenadas de subsecuencia, si fuera necesario, y se designan parámetros de programa de algoritmo de secuencias. Pueden usarse parámetros de programa por defecto, o pueden designarse parámetros alternativos. El algoritmo de comparación de secuencias calcula entonces el porcentaje de identidades de secuencia para las secuencias de prueba con respecto a la secuencia de referencia,

basándose en los parámetros de programa.

Una "ventana de comparación", como se usa en el presente documento, incluye referencia a un segmento de una cualquiera del número de posiciones contiguas de aproximadamente 10 a aproximadamente 500, aproximadamente 25 a aproximadamente 200, 50 a aproximadamente 150, en la que una secuencia puede compararse con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias se alineen óptimamente. Métodos de alineamiento de secuencias para comparación son muy conocidos en la técnica. El alineamiento óptimo de secuencias para comparación puede realizarse, por ejemplo, por el algoritmo de homología local de Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), por el algoritmo de alineamiento por homología de Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), por la búsqueda del método de similitud de Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), por implementaciones computerizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), o por alineamiento manual e inspección visual (véase, por ejemplo, *Current Protocols in Molecular Biology* (Ausubel et al., eds. 1995 suplemento)).

Un ejemplo de un algoritmo útil es PILEUP. PILEUP crea un alineamiento múltiple de secuencias de un grupo de secuencias relacionadas usando alineamientos por pares progresivos para mostrar relación y porcentaje de identidad de secuencias. También representa un árbol o dendograma que muestra las relaciones de agrupación usadas para crear el alineamiento. PILEUP usa una simplificación del método de alineamiento progresivo de Feng & Doopoco, *J. Mol. Evol.* 35:351-360 (1987). El método usado es similar al método descrito por Higgins & Sharp, *CABIOS* 5:151-153 (1989). El programa puede alinear hasta 300 secuencias, cada una de una longitud máxima de 5.000 nucleótidos o aminoácidos. El procedimiento de alineamiento múltiple empieza con el alineamiento por parejas de las dos secuencias más similares, produciendo una agrupación de dos secuencias alineadas. Esta agrupación se alinea entonces con la siguiente secuencia más relacionada o agrupación de secuencias alineadas. Dos agrupaciones de secuencias están alineadas por una simple extensión del alineamiento por pares de dos secuencias individuales. El alineamiento final se logra por una serie de alineamientos por pares progresivo. El programa se ejecuta designando secuencias específicas y sus coordenadas de aminoácidos o nucleótidos para regiones de comparación de secuencias y designando los parámetros de programa. Usando PILEUP, una secuencia de referencia se compara con otras secuencias de prueba para determinar la relación del porcentaje de identidad de secuencias usando los siguientes parámetros: peso por hueco por defecto (3,00), peso por longitud de hueco por defecto (0,10) y huecos de extremos ponderados. PILEUP puede obtenerse del paquete de software de análisis de secuencias GCG, por ejemplo, versión 7.0 (Devereaux et al., *Nuc. Acids Res.* 12:387-395 (1984)).

Otro ejemplo de algoritmo que es adecuado para determinar el porcentaje de identidad de secuencias y similitud de secuencias son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul et al., *Nuc. Acids Res.* 25:3389-3402 (1977) y Altschul et al., *J. Mol. Biol.* 215:403-410 (1990), respectivamente. El software para realizar los análisis de BLAST está públicamente disponible del Centro Nacional para Información Biotecnológica (<http://www.ncbi.nlm.nih.gov/>). Este algoritmo implica primero identificar pares de secuencias de alta puntuación (HSP) identificando palabras cortas de longitud W en la secuencia de búsqueda, que tanto coinciden como cumplen alguna puntuación umbral de valor positivo T cuando se alinean con una palabra de la misma longitud en una secuencia de base de datos. T se denomina el umbral de puntuación de la palabra de vecindad (Altschul et al., arriba). Estos éxitos de palabra de vecindad inicial actúan de semillas para iniciar búsquedas para encontrar HSP más largas que las contienen. Los éxitos de palabra se extienden en ambas direcciones a lo largo de cada secuencia tan largo como las puntuaciones de alineamiento acumulado puedan aumentarse. Las puntuaciones acumuladas se calculan usando, para secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de residuos de correspondencia; siempre > 0) y N (puntuación de penalización para residuos no de correspondencia; siempre < 0). Para secuencias de aminoácidos se usa una matriz de puntuación para calcular la puntuación acumulada. Los éxitos de extensión de la palabra en cada dirección se detienen cuando: la puntuación de alineamiento acumulada disminuye la cantidad X de su máximo valor logrado; la puntuación acumulada tiende a cero o por debajo, debido a la acumulación de uno o más alineamientos de residuos de puntuación negativa; o se alcanza el extremo de cualquier secuencia. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad del alineamiento. El programa BLASTN (para secuencias de nucleótidos) usa por defecto una longitud de palabra (W) de 11, una esperanza (E) o 10, M=5, N=-4 y una comparación de ambas cadenas. Para secuencias de aminoácidos, el programa BLASTP usa por defecto una longitud de palabra de 3, y esperanza (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alineamientos (B) de 50, esperanza (E) de 10, M=5, N=-4, y una comparación de ambas hebras.

El algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). Una medida de similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que proporciona una indicación de la probabilidad por la que una coincidencia entre dos secuencias de nucleótidos o de aminoácidos se produciría por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación del ácido nucleico de prueba con el ácido nucleico de referencia es inferior a aproximadamente 0,2, más preferentemente inferior a aproximadamente 0,01, y lo más preferentemente inferior a aproximadamente 0,001.

III. Composiciones de la presente invención

La invención proporciona composiciones que comprenden vectores adenovirales quiméricos. En algunas realizaciones, los vectores adenovirales quiméricos de la invención comprenden un primer promotor unido operativamente a un ácido nucleico que codifica un polipéptido heterólogo y un segundo promotor unido operativamente a un ácido nucleico que codifica un agonista de TLR-3. El primer y segundo promotores pueden ser iguales o diferentes. En algunas realizaciones, el primer y segundo promotores están seleccionados independientemente de: el promotor de beta-actina y un promotor del CMV.

En algunos aspectos de la invención, el vector adenoviral quimérico comprende el genoma adenoviral (sin los genes E1 y E3) y un ácido nucleico que codifica un gen que activa IRF-3 y otras moléculas de señalización en la dirección 3' de TLR-3. El vector quimérico puede administrarse a una célula que expresa el gen E1 de Ad de forma que el adenovirus recombinante (rAd) se produzca por la célula. Este rAd puede recogerse y es capaz de una única ronda de infección que proporcionará la composición transgénica a otra célula dentro de un mamífero con el fin de provocar respuestas inmunitarias al polipéptido heterólogo.

A. Vectores adenovirales adecuados

En algunas realizaciones, el vector adenoviral es el adenovirus 5, que incluye, por ejemplo, Ad5 con deleciones de las regiones E1/E3 y Ad5 con una deleción de la región E4. Otros vectores adenovirales adecuados incluyen las cepas 2, cepas 4 y 7 probadas por vía oral, adenovirus 40 y 41 entéricos y otras cepas (por ejemplo, Ad34) que son suficientes para administrar un antígeno y que provocan una respuesta inmunitaria adaptativa al antígeno de transgén [Lubeck et al., Proc Natl Acad Sci U S A, 86(17), 6763-6767 (1989); Shen et al., J Virol, 75(9), 4297-4307 (2001); Bailey et al., Virology, 202(2), 695-706 (1994)]. En algunas realizaciones, el vector adenoviral es un vector adenoviral vivo incompetente para la replicación (tal como rAd5 con E1 y E3 delecionadas), vector adenoviral vivo y atenuado (tal como los virus de deleción de E1B55K), o un vector adenoviral vivo con replicación no mutante.

Las secuencias de control de la transcripción y de la traducción en vectores de expresión que van a usarse en células de vertebrado transformantes *in vivo* pueden proporcionarse por fuentes virales. Por ejemplo, promotores y potenciadores comúnmente usados se derivan, por ejemplo, de beta-actina, adenovirus, virus simio (SV40) y citomegalovirus humano (CMV). Por ejemplo, son adecuados vectores que permiten la expresión de proteínas bajo la dirección de un promotor del CMV, promotor temprano del SV40, promotor tardío del SV40, promotor de metalotioneína, promotor del virus del tumor mamario murino, promotor del virus del sarcoma de Rous, promotor de transductor u otros promotores que muestran ser eficaces para la expresión en células de mamífero. Pueden usarse secuencias de promotor genómico viral, de control y/o señal adicionales, siempre que tales secuencias de control sean compatibles con la célula huésped elegida.

B. Polipéptidos heterólogos

Ácidos nucleicos que codifican polipéptidos heterólogos adecuados pueden derivarse de antígenos, tales como, por ejemplo, antígenos virales, antígenos bacterianos, antígenos de cáncer, antígenos fúngicos o antígenos de parásito.

Los antígenos virales pueden derivarse de, por ejemplo, virus de la inmunodeficiencia humana (por ejemplo, gag (p55 y p160), pol, env (gp120 y gp41) como se exponen en Shiver et al. Nature 415(6869):331 (2002); las secuencias genómicas del VIH expuestas en los N.º de acceso de Genbank EF363127; EF363126; EF363125; EF363124; EF363123; EF363122; EF192592; y EF192591; las secuencias de gag del VIH expuestas en los N.º de acceso de Genbank EF396891; EF396890; EF396889; EF396888; EF396887; EF396886; EF396885; EF396884; EF396883; EF396882; EF396881; EF396880; EF396879; EF396878; EF396877; EF396876; EF396875; EF396874; EF396873; y EF396872; las secuencias de pol del VIH expuestas en los N.º de acceso de Genbank EF396810; EF396809; EF396808; EF396807; EF396806; EF396805; EF396804; EF396803; EF396802; EF396801; EF396800; EF396799; EF396798; EF396797; EF396796; EF396795; EF396794; EF396793; EF396792; y EF396791; y las secuencias de env del VIH expuestas en los N.º de acceso de Genbank 9: EF367234; EF367233; EF367232; EF367231; EF367230; EF367229; EF367228; EF367227; EF367226; EF367225; EF367224; y EF367223, virus del papiloma humano (por ejemplo, proteína de la cápsida L1 como se describe en, por ejemplo, Donnelly et al. J Infect Dis. 173:314 (1996) y las secuencias expuestas en los N.º de acceso de Genbank EF362755; EF362754; NC_001694; NC_001693; NC_001691; NC_001690; NC_005134; NC_001458; NC_001457; NC_001354; NC_001352; NC_001526; y X94164), virus de Epstein Barr, virus del herpes simple, virus del herpes humano, rinovirus, virus de Coxsackie, enterovirus, hepatitis A, B, C y E (por ejemplo, antígeno de superficie de la hepatitis B como se describe en, por ejemplo, Lubeck et al, PNAS USA 86:6763 (1989) y las secuencias expuestas en los N.º de acceso de GenBank AB236481; AB236471; AB206501; AB206489; AB206487; AB221788; AB221777; AB221773; AR933671; AR933670; AB236514; AB236513; AB236512; AB236511; AB236510; AB236509; AB236508; AB236507); NS5 de la hepatitis C (véanse, por ejemplo, los N.º de acceso de Genbank X59609; DQ911563; S71627; S70787; S70786; S70341; S62220; S70790; S70789; S70788; y AB204642)), virus de las paperas, virus de la rubeola, virus del sarampión, virus de la polio, virus de la viruela, virus de la rabia y virus de la varicela zóster.

Antígenos de la gripe incluyen, por ejemplo, hemaglutinina (HA), proteína de la matriz 1 (M1) y nucleoproteína (NP) (véase, por ejemplo, Donnelly, et al, Vaccine 15:865 (1997) y las secuencias de HA de la gripe expuestas en los N.º de acceso de Genbank AB294219; AB294217; AB294215; AB294213; EF102944; EF102943; EF102942; EF102941; EF102940; EF102939; EF102938; EF102937; EF102936; EF102935; EF102934; EF102933; DQ643982; DQ464354; 5 CY019432; CY019424; CY019416; CY019408; CY019400; CY019392; CY019384; CY019376; CY019368; CY019360; CY019352; EF124794; EF110519; EF110518; EF165066; EF165065; EF165064; y EF165063; las secuencias de M1 de la gripe expuestas en los N.º de acceso de Genbank AB292791; CY019980; CY019972; CY019964; CY019956; CY019948; CY019940; CY019628; CY019652; CY019644; CY019932; CY019924; 10 CY019916; CY019908; CY019900; CY019892; CY019884; CY019876; CY019868; CY019860; y las secuencias de NP de la gripe expuestas en los N.º de acceso de Genbank AB292790; CY019461; CY019974; CY019966; CY019958; CY019950; CY019942; CY019630; CY019654; CY019646; CY019934; CY019926; CY019918 CY019910; CY019902; CY019894; CY019886; CY019878; CY019870; y CY019862.

Antígenos virales adecuados también incluyen, por ejemplo, proteínas no estructurales virales. El término “proteína no estructural viral”, como se usa en el presente documento, se refiere a proteínas codificadas por ácido nucleico viral que no codifican polipéptidos estructurales, tales como aquellos que hacen la cápside o la proteína alrededor de un virus. Proteínas no estructurales incluyen aquellas proteínas que promueven la replicación de ácido nucleico viral y expresión génica viral tales como, por ejemplo, las proteínas no estructurales 1, 2, 3 y 4 (NS1, NS2, NS3 y NS4, respectivamente) de la encefalitis equina venezolana (VEE), EEE, o virus del bosque de Semliki [Dubensky et al., J 15 Virol, 70(1), 508-519 (1996); Petrakova et al J Virol 2005 79(12): 7597-608; patentes de EE.UU. N.º 5.185.440; 5.739.026; 6.566.093; y 5.814.482. Varios ejemplos representativos de alfavirus adecuados incluyen Aura (ATCC VR-368), virus de Bebaru (ATCC VR-600, ATCC VR-1240), Cabassou (N.º de acceso de Genbank AF398387, ATCC VR-922), virus de Chikungunya (ATCC VR-64, ATCC VR-1241), virus de la encefalitis equina oriental (N.º de acceso de Genbank AY705241, AY705240, ATCC VR-65, ATCC VR-1242), Fort Morgan (ATCC VR-924), virus de Getah 20 (ATCC VR-369, ATCC VR-1243), Kyzilagach (ATCC VR-927), Mayaro (ATCC VR-66), virus de Mayaro (ATCC VR-1277), Middleburg (ATCC VR-370), virus de Mucambo (ATCC VR-580, ATCC VR-1244), Ndumu (ATCC VR-371), virus de Pixuna (ATCC VR-372, ATCC VR-1245), virus del río Ross (ATCC VR-373, ATCC VR-1246), bosque de Semliki (N.º de acceso de Genbank AJ251359, ATCC VR-67, ATCC VR-1247), virus de Sindbis (N.º de acceso de Genbank J02363, ATCC VR-68, ATCC VR-1248), Tonate (ATCC VR-925), Trinititi (ATCC VR-469), Una (ATCC VR-374), encefalomiелitis equina venezolana (ATCC VR-69), virus de la encefalomiелitis equina venezolana (N.º de acceso de Genbank AY986475, AY973944, NC 001449, ATCC VR-923, ATCC VR-1250 ATCC VR-1249, ATCC VR-532), encefalomiелitis equina occidental (ATCC VR-70, ATCC VR-1251, ATCC VR-622, ATCC VR-1252), Whataroa 25 (ATCC VR-926), y Y-62-33 (ATCC VR-375).

Antígenos bacterianos pueden derivarse de, por ejemplo, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Helicobacter pylori*, *Streptococcus bovis*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Listeria monocytogenes*, *Mycobacterium tuberculosis*, *Mycobacterium leprae*, *Corynebacterium diphtheriae*, *Borrelia burgdorferi*, *Bacillus anthracis*, *Bacillus cereus*, *Clostridium botulinum*, *Clostridium difficile*, *Salmonella typhi*, *Vibrio cholerae*, *Haemophilus influenzae*, *Bordetella pertussis*, *Yersinia pestis*, *Neisseria gonorrhoeae*, *Treponema pallidum*, *Mycoplasma sp.*, *Neisseria ransducker s*, *Legionella pneumophila*, *Rickettsia typhi*, *Chlamydia trachomatis* y *Shigella dysenteriae*, *Vibrio cholera* (por ejemplo, toxina del cólera subunidad B como se expone en los N.º de acceso de Genbank U25679; A09803; EF158842; X76391; AF390572; pelo co-regulado por la toxina del cólera 35 (TCP) como se describe en Wu et al., Infection and Immunity Vol. 69(12):7695 (2001) y se expone en los N.º de acceso de Genbank NC_002505 y AE004169); *Helicobacter pylori* (VacA como se expone en los N.º de acceso de Genbank AY848858; AF042737; AF042736; AF042735; AF042734; NC_000921; CagA como se expone en los N.º de acceso de Genbank AF043490; AF043489; AF043488; AF043487; NAP como se expone en los N.º de acceso de Genbank AF284121; AF284120; AF284119; AF284118; AF284117; AF284116; AB045143; AB045142; AF227081; AF227080; AF227079; AF227078; AF227077; AF227076; AF227075; AF227074; Hsp o catalasa como se expone en los N.º de acceso de Genbank NC_000921; ureasa como se expone en los N.º de acceso de Genbank AM417610; 40 AM417609; AM417608; AM417607; AM417606; AM417605; AM417604; AM417603; AM417602; AM417601; y AM417600; antígenos de *E. coli* como se exponen en los N.º de acceso de Genbank NC_000913; U00096; NC_002655; BA000007; AE014075; incluyendo antígenos fimbriales de *E. coli* como se exponen en los N.º de acceso de Genbank AB214865; AB214864; AB214863; AB214862; enterotoxina lábil al calor de *E. coli* como se expone en los N.º de acceso de Genbank X83966; V00275; X83966; J01646; V00275; M35581; M17873; M17874; 55 K01995; M61015; M17894; M17101; K00433.

Los antígenos de parásito pueden derivarse de, por ejemplo, *Giardia lamblia*, *Leishmania sp.*, *Trypanosoma sp.*, *Trichomonas sp.*, *Plasmodium sp.* (por ejemplo, antígenos de proteína de la superficie de *P. falciparum* tales como secuencias de pfs25 como se exponen en los N.º de acceso de Genbank XM_001347551; X07802; AF193769; 60 AF179423; AF154117; y AF030628, secuencias de pfs28 como se exponen en los N.º de acceso de GenBank L25843, secuencias de pfs45 como se exponen en los N.º de acceso de Genbank EF158081; EF158079; EF158078; EF158076; EF158075; y EF158085, secuencias de pfs84, pfs 48/45 como se exponen en los N.º de acceso de Genbank AF356146; AF356145; AF356144; AF356143; AF356142; AF356141; AF356140; AF356139; AF356138; AF356137; AF356136; AF356135; AF356134; AF356133; AF356132; AF356131; AF356130; AF356129; AF356128; 65 AF356127, secuencias de pfs 230 como se exponen en los N.º de acceso de Genbank NC_000910; XM_001349564; AE001393; L22219; L08135; y AF269242, antígenos de *P. vivax* tales como secuencias de Pvs25 como se exponen

en los N.º de acceso de Genbank DQ641509; DQ641508; DQ641507; AY639972; AY639971; AY639970; AY639969; AY639968; AY639967; AY639966; y AY639965; y secuencias de Pvs28 como se exponen en los N.º de acceso de Genbank AB033364; AB033363; AB033362; AB033361; AB033360; AB033359; AB033358; AB033357; AB033356; B033355; AB033354; AB033353; AB033352; AB033351; AB033350; AB033349; AB033348; AB033347; AB033346; y AB033345), *Schistosoma sp.*, *Mycobacterium tuberculosis* (por ejemplo, secuencias de Ag85 como se exponen en los N.º de acceso de Genbank AX253506; AX253504; AX253502; y AX211309; MPT64, ESAT-6, CFP10, R8307, MTB-32 MTB-39, CSP, LSA-1, LSA-3, EXP1, SSP-2, SALSA, STARP, GLURP, MSP-1, MSP-2, MSP-3, MSP-4, MSP-5, MSP-8, MSP-9, AMA-1, proteína de la membrana integral tipo 1, secuencias de RESA, EBA-175 y DBA como se exponen en los N.º de acceso de Genbank BX842572; BX842573; BX842574; BX842575; BX842576; BX842577; BX842578; BX842579; BX842580; BX842581; BX842582; BX842583; BX842584 y NC_000962, secuencias de HSP65 como se exponen en los N.º de acceso de Genbank AY299175; AY299174; AY299144; AF547886; y AF547885).

Los antígenos de cáncer incluyen, por ejemplo, antígenos expresados, por ejemplo, en cáncer de colon, cáncer de estómago, cáncer pancreático, cáncer de pulmón, cáncer de ovario, cáncer de próstata, cáncer de mama, cáncer de piel (por ejemplo, melanoma), leucemia, linfoma o mieloma, antígenos de cáncer a modo de ejemplo incluyen, por ejemplo, HPV L1, HPV L2, HPV E1, HPV E2, fosfatasa alcalina placentaria, AFP, BRCA1, Her2/neu, CA 15-3, CA 19-9, CA-125, CEA, Hcg, activador del plasminógeno tipo urocinasa (Upa), inhibidor del activador del plasminógeno.

Los antígenos fúngicos pueden derivarse de, por ejemplo, *Tinea pedis*, *Tinea corporis*, *Tinea cruris*, *Tinea unguium*, *Cladosporium carionii*, *Coccidioides immitis*, *Candida sp.*, *Aspergillus fumigatus* y *Pneumocystis carinii*.

Los ácidos nucleicos que codifican polipéptidos inmunogénicos se producen normalmente por métodos de ADN recombinante (véase, por ejemplo, Ausubel, et al. ed. (2001) Current Protocols in Molecular Biology). Por ejemplo, las secuencias de ADN que codifican el polipéptido inmunogénico pueden ensamblarse de fragmentos de ADNc y conectores de oligonucleótido cortos, o de una serie de oligonucleótidos, o amplificarse de ADNc usando cebadores apropiados para proporcionar un gen sintético que es capaz de insertarse en un vector de expresión recombinante (es decir, un vector plasmídico o un vector viral) y expresarse en una unidad transcripcional recombinante. Una vez se produce el ácido nucleico que codifica un polipéptido inmunogénico, puede insertarse en un vector de expresión recombinante que es adecuado para la expresión *in vivo* o *ex vivo*.

Los vectores de expresión recombinante contienen una secuencia de ADN que codifica un polipéptido inmunogénico unido operativamente a elementos reguladores de la transcripción o de la traducción adecuados derivados de genes de mamífero o virales. Tales elementos reguladores incluyen un promotor de la transcripción, una secuencia de operador opcional para controlar la transcripción, una secuencia que codifica sitios de unión al ribosoma de ARNm adecuados, y secuencias que controlan la terminación de la transcripción y traducción. Adicionalmente puede incorporarse un origen de replicación y un marcador de selección para facilitar el reconocimiento de transformantes. Los genes utilizados en los vectores de expresión recombinantes pueden dividirse entre más de un virus de forma que los productos génicos estén en dos vectores diferentes, y los vectores se usan para la co-transducción para proporcionar todos los productos génicos en trans. Puede haber motivos para dividir los productos génicos tales como limitaciones de tamaño para inserciones, o toxicidad de los productos génicos combinados a las líneas celulares productoras de virus.

C. Agonistas de TLR-3

En la presente invención, un agonista de TLR-3 que es un ARNbc heterólogo se usa para estimular el reconocimiento inmunitario de un antígeno de interés. Los agonistas de TLR-3 incluyen, por ejemplo, ARN de horquilla corta, ARN viralmente derivado, segmentos cortos de ARN que pueden formar hebras dobles o ARN de horquilla corta, y ARN interferente pequeño (ARNip). En una realización de la invención, el agonista de TLR-3 es ARNbc viralmente derivado, tal como, por ejemplo, un ARNbc derivado de un virus de Sindbis o productos intermedios de ARNbc viral [Alexopoulou et al, Nature 413:732-8 (2001)]. En algunas realizaciones, el agonista de TLR-3 es una ARN de horquilla corta. Las secuencias de ARN de horquilla corta normalmente comprenden dos secuencias complementarias unidas por una secuencia de conector. La secuencia de conector particular no es un aspecto crítico de la invención. Puede usarse cualquier secuencia de conector apropiada, mientras que no interfiera con la unión de las dos secuencias complementarias para formar un ARNbc.

En algunas realizaciones, el ARN de horquilla corta comprende una secuencia expuesta en SEQ ID NOS: 3, 4, 5, 8, 9, 10, 11 o 12, una secuencia con identidad sustancial con una secuencia expuesta en SEQ ID NOS: 3, 4, 5, 8, 9, 10, 11 o 12, o una variante de una secuencia expuesta en SEQ ID NOS: 3, 4, 5, 8, 9, 10, 11 o 12. En ciertas realizaciones, el ARNbc que es un agonista de TLR-3 no codifica un polipéptido particular, pero produce una citocina pro-inflamatoria (por ejemplo, IL-6, IL-8, TNF-alfa, IFN-alfa, IFN-beta) cuando se pone en contacto con una célula respondedora (por ejemplo, una célula dendrítica, una célula mononuclear de sangre periférica o un macrófago) *in vitro* o *in vivo*. En algunos casos, el ácido nucleico que codifica el agonista de TLR-3 (un ARNbc expresado) y el vector adenoviral quimérico que comprende un ácido nucleico que codifica un antígeno heterólogo se administran en la misma formulación. En algunas realizaciones, el ácido nucleico que codifica el agonista de TLR-3 y el ácido nucleico que codifica el antígeno heterólogo están bajo el control de diferentes promotores.

Están comercialmente disponibles varios análogos químicamente sintetizados para ARN bicatenario. Éstos incluyen poli-inosina-ácido policitídílico (poli I:C), ácido poliadenílico-poliuridílico (poli A:U), y poli I:poli C. Anticuerpos (o reticulación de anticuerpos) para TLR-3 también pueden conducir a la producción de IFN-beta o citocinas pro-inflamatorias [Matsumoto et al, *Biochem. Biophys. Res. Commun.* 24:1364 (2002), de Bouteiller et al, *J Biol. Chem.* 18:38133-45 (2005)]. También pueden obtenerse segmentos de ARN comercialmente disponibles de cualquier secuencia mediante fuentes tales como Invitrogen.

IV. Composiciones farmacéuticas

Las composiciones farmacéuticas que comprenden los vectores descritos en el presente documento también pueden contener otros compuestos, que pueden ser biológicamente activos o inactivos. Los polipéptidos pueden, pero no necesitan, conjugarse con otras macromoléculas como se describen, por ejemplo, en las patentes de EE.UU. N.º 4.372.945 y 4.474.757. Las composiciones farmacéuticas pueden usarse generalmente para fines profilácticos y terapéuticos. Las composiciones farmacéuticas pueden estar compuestas de métodos para proteger contra la degradación del estómago de forma que el vector adenoviral quimérico administrado pueda llegar a las localizaciones deseadas. Para el entorno oral, varias de éstas están disponibles, que incluyen los sistemas de liberación Eudragit y TimeClock, además de otros métodos específicamente diseñados para adenovirus [Lubeck et al., *Proc Natl Acad Sci U S A*, 86(17), 6763-6767 (1989); Choraasia y Jain, *J Pharm Pharm Sci*, 6(1), 33-66 (2003)]. También se han descrito ya varios métodos para la microencapsulación de ADN y fármacos para administración oral (véase, por ejemplo, la publicación de patente de EE.UU. N.º 2004043952). En algunas realizaciones, el sistema Eudragit se usará para administrar el vector adenoviral quimérico al intestino delgado inferior. Sin embargo, también debería funcionar la administración a otras localizaciones del intestino delgado.

Como se observa anteriormente, los vectores adenovirales quiméricos de la invención pueden administrarse usando cualquier sistema de administración conocido para aquellos expertos habituales en la materia. Numerosas técnicas de administración génica son muy conocidas en la técnica, tales como aquellas descritas por Rolland (1998) *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, y referencias citadas en su interior.

Será evidente que una composición inmunogénica puede contener sales farmacéuticamente aceptables de los polinucleótidos que codifican los polipéptidos heterólogos (por ejemplo, polipéptidos inmunogénicos). Tales sales pueden prepararse a partir de bases no tóxicas farmacéuticamente aceptables, que incluyen bases orgánicas (por ejemplo, sales de aminas primarias, secundarias y terciarias y aminoácidos básicos) y bases inorgánicas (por ejemplo, sales de sodio, potasio, litio, amonio, calcio y magnesio). Algunos ejemplos de sales particulares incluyen solución salina tamponada con fosfato y solución salina para inyección.

Puede emplearse cualquier vehículo adecuado conocido para aquellos expertos habituales en la materia en las composiciones farmacéuticas de la presente invención. Vehículos adecuados incluyen, por ejemplo, agua, solución salina, alcohol, una grasa, una cera, un tampón, un vehículo sólido, tal como manitol, lactosa, almidón, estearato de magnesio, sacarina sódica, talco, celulosa, glucosa, sacarosa y carbonato de magnesio, o microesferas biodegradables (por ejemplo, polilactato-poliglicolato). Microesferas biodegradables adecuadas se desvelan, por ejemplo, en las patentes de EE.UU. N.º 4.897.268; 5.075.109; 5.928.647; 5.811.128; 5.820.883. El polipéptido inmunogénico y/o virus portador pueden encapsularse dentro de la microesfera biodegradable o asociarse a la superficie de la microesfera.

Tales composiciones también pueden comprender tampones (por ejemplo, solución salina tamponada neutra o solución salina tamponada con fosfato), hidratos de carbono (por ejemplo, glucosa, manosa, sacarosa o dextranos), manitol, proteínas, polipéptidos o aminoácidos tales como glicina, antioxidantes, bacteriostáticos, agentes quelantes tales como EDTA o glutatión, adyuvantes (por ejemplo, hidróxido de aluminio), solutos que convierten la formulación en isotónica, hipotónica o débilmente hipertónica con la sangre de un receptor, agentes de suspensión, espesantes y/o conservantes. Alternativamente, las composiciones de la presente invención pueden formularse como un liofilizado. Los compuestos también pueden encapsularse dentro de liposomas usando tecnología muy conocida.

En algunas divulgaciones, las composiciones comprenden además un adyuvante. Adyuvantes adecuados incluyen, por ejemplo, los lípidos y compuestos no de lípido, toxina del cólera (CT), subunidad B de CT, derivado de CT CTK63, enterotoxina lábil al calor de *E. coli* (LT), derivado de LT LTK63, Al(OH)₃ y ácidos orgánicos poliónicos como se describen en, por ejemplo, el documento WO 04/020592, Anderson y Crowle, *Infect. Immun.* 31(1):413-418 (1981), Roterman et al, *J. Physiol. Pharmacol.* 44(3):213-32 (1993), Arora y Crowle, *J. Reticuloendothel.* 24(3):271-86 (1978), y Crowle y May, *Infect. Immun.* 38(3):932-7 (1982)). Ácidos orgánicos poliónicos adecuados incluyen, por ejemplo, ácido 6,6'-[3,3'-dimetil[1,1'-bifenil]-4,4'-diil]bis(azo)bis[4-amino-5-hidroxi-1,3-naftaleno-disulfónico] (azul de Evans) y ácido 3,3'-[1,1'-bifenil]-4,4'-diilbis(azo)bis[4-amino-1-naftaleno-sulfónico] (rojo Congo). Se apreciará por aquellos expertos en la materia que los ácidos orgánicos poliónicos pueden usarse para cualquier método de vacunación genética conjuntamente con cualquier tipo de administración.

Otros adyuvantes adecuados incluyen inmunomoduladores tópicos tales como miembros de la familia de las imidazoquinolinas tales como, por ejemplo, imiquimod y resiquimod (véase, por ejemplo, Hengge et al, *Lancet Infect. Dis.* 1(3):189-98 (2001)). También podrían usarse agonistas de TLR-3 expresados (por ejemplo, ARNbc) y agonistas

de TLR-7 (por ejemplo, ARNm).

Adyuvantes adecuados adicionales están comercialmente disponibles como, por ejemplo, adyuvantes basados en alumbre adicionales (por ejemplo, Alhydrogel, Rehydralgel, fosfato de aluminio, Algammulin); adyuvantes basados en aceite (adyuvante incompleto de Freund y adyuvante completo (Difco Laboratories, Detroit, Mich.), Specol, RIBI, TiterMax, Montanide ISA50 o Seppic MONTANIDE ISA 720); adyuvantes basados en copolímeros de bloque no iónicos, citocinas (por ejemplo, GM-CSF o ligando de Flat3); adyuvante 65 de Merck (Merck and Company, Inc., Rahway, N. J.); AS-2 (SmithKline Beecham, Philadelphia, Pa.); sales de calcio, hierro o cinc; una suspensión insoluble de tirosina acilada; azúcares acilados; polisacáridos catiónicamente o aniónicamente derivatizados; polifosfacenos; microesferas biodegradables; monofosforil lípido A y Quil A. Citocinas, tales como GM-CSF o interleucina-2, -7, o -12, también son adyuvantes adecuados. También pueden usarse hemocianinas (por ejemplo, hemocianina de lapa californiana) y hemoeritrinas en la invención. Adyuvantes de polisacárido tales como, por ejemplo, quitina, quitosano y quitina desacetilada también son adecuados como adyuvantes. Otros adyuvantes adecuados incluyen muramil dipéptido (MDP, N-acetilmuramil-L-alanil-D-isoglutamina), peptidoglicanos bacterianos y sus derivados (por ejemplo, treonil-MDP y MTPPE). También puede usarse BCG y esqueleto de la pared celular de BCG (CWS) como adyuvantes, con o sin dimicolato de trehalosa. El dimicolato de trehalosa puede usarse él mismo (véase, por ejemplo, la patente de EE.UU. N.º 4.579.945). También son útiles endotoxinas desintoxicadas como adyuvantes solos o en combinación con otros adyuvantes (véanse, por ejemplo, las patentes de EE.UU. N.º 4.866.034; 4.435.386; 4.505.899; 4.436.727; 4.436.728; 4.505.900; y 4.520.019. Las saponinas QS21, QS17, QS7 también son útiles como adyuvantes (véanse, por ejemplo, la patente de EE.UU. N.º 5.057.540; documentos EP 0362 279; WO 96/33739; y WO 96/11711). Otros adyuvantes adecuados incluyen Montanide ISA 720 (Seppic, Francia), SAF (Chiron, Calif., Estados Unidos), ISCOEM (CSL), MF-59 (Chiron), la serie SBAS de adyuvantes (por ejemplo, SBAS-2, SBAS-4 o SBAS-6 o variantes de los mismos, disponibles de SmithKline Beecham, Rixensart, Bélgica), Detox (Corixa, Hamilton, Mont.) y RC-529 (Corixa, Hamilton, Mont).

También se contemplan superantígenos para su uso como adyuvantes en la presente divulgación. Los superantígenos incluyen exoproteínas de *Staphylococcus*, tales como las enterotoxinas α , β , γ y Δ de *S. aureus* y *S. epidermidis*, y las exotoxinas de *E. coli* α , β , γ y Δ . Enterotoxinas de *Staphylococcus* comunes se conocen como enterotoxina estafilocócica A (SEA) y enterotoxina estafilocócica B (SEB), describiéndose las enterotoxinas hasta E (SEE) (Rott et al., 1992). *Streptococcus pyogenes* B (SEB), enterotoxina de *Clostridium perfringens* (Bowness et al., 1992), proteína asociada a la membrana citoplásmica (CAP) de *S. pyogenes* (Sato et al., 1994) y toxina 1 de síndrome de choque tóxico (TSST 1) de *S. aureus* (Schwab et al., 1993) son superantígenos útiles adicionales.

Dentro de las composiciones farmacéuticas proporcionadas en el presente documento, la composición de adyuvante puede diseñarse para inducir, por ejemplo, una respuesta inmunitaria predominantemente del tipo Th1 o Th2. Altos niveles de citocinas de tipo Th1 (por ejemplo, IFN-gamma, TNF-alfa, IL-2 e IL-12) tienden a favorecer la inducción de respuestas inmunitarias celulares a un antígeno administrado. A diferencia, altos niveles de citocinas tipo Th2 (por ejemplo, IL-4, IL-5, IL-6 e IL-10) tienden a favorecer la inducción de respuestas inmunitarias humorales. Tras la administración oral de una composición que comprende un polipéptido inmunogénico como se ha proporcionado en el presente documento, normalmente se provocará una respuesta inmunitaria que incluye respuestas del tipo Th1 y Th2.

Las composiciones descritas en el presente documento pueden administrarse como parte de una formulación de liberación sostenida (es decir, una formulación tal como una cápsula o esponja que efectúa una liberación lenta de compuesto tras la administración). Tales formulaciones pueden prepararse generalmente usando tecnología muy conocida (véase, por ejemplo, Coombes et al. (1996) Vaccine 14:1429-1438). Las formulaciones de liberación sostenida pueden contener un polipéptido, polinucleótido o anticuerpo dispersado en una matriz portadora y/o contenido dentro de un depósito rodeado por una membrana de control de la velocidad.

Los vehículos para su uso dentro de tales formulaciones son biocompatibles, y también pueden ser biodegradables; preferentemente, la formulación proporciona un nivel relativamente constante de liberación de componente activo. Tales vehículos incluyen micropartículas de poli(lactida-co-glicolida), además de poliacrilato, látex, almidón, celulosa y dextrano. Otros vehículos de liberación retardada incluyen biovectores supramoleculares, que comprenden un núcleo hidrófilo no líquido (por ejemplo, un polisacárido u oligosacárido reticulado) y, opcionalmente, una capa externa que comprende un compuesto anfílico (véanse, por ejemplo, los documentos WO 94/20078; WO 94/23701; y WO 96/06638). La cantidad de compuesto activo contenido dentro de una formulación de liberación sostenida depende del sitio de implantación, la velocidad y duración esperada de la liberación, y la naturaleza de la afección que va a tratarse o prevenirse.

Las composiciones farmacéuticas pueden presentarse en recipientes de dosis unitaria o de múltiples dosis, tales como ampollas o viales sellados. Tales recipientes están preferentemente herméticamente sellados para preservar la esterilidad de la formulación hasta su uso. En general, las formulaciones pueden almacenarse como suspensiones, disoluciones o emulsiones en vehículos aceitosos o acuosos. Alternativamente, una composición farmacéutica puede almacenarse en una condición liofilizada que requiere solo la adición de un vehículo líquido estéril inmediatamente antes de uso.

V. Usos terapéuticos de la invención

Un aspecto de la presente invención implica las composiciones inmunogénicas para su uso en un método de tratamiento que provoca una respuesta inmunitaria específica de antígeno de un sujeto o paciente con una enfermedad tal como, por ejemplo, una infección viral, infección bacteriana, una infección parasítica, una infección fúngica, o cáncer. Como se usa en el presente documento, un "sujeto" o un "paciente" se refiere a cualquier animal de sangre caliente, tal como, por ejemplo, un roedor, un felino, un canino o un primate, preferentemente un ser humano. Las composiciones inmunogénicas pueden usarse para tratar cualquier etapa de la enfermedad, es decir, en las etapas de pre-cáncer, cáncer o metastásicas, o para prevenir la enfermedad. Por ejemplo, las composiciones descritas en el presente documento pueden usarse para tratar una enfermedad viral tal como VIH o hepatitis o para la prevención o tratamiento de cáncer. Dentro de tales métodos, las composiciones farmacéuticas normalmente se administran a un paciente. El paciente puede o puede no estar afectado por la enfermedad o trastorno (por ejemplo, una infección viral, una infección bacteriana, o cáncer). Por consiguiente, las composiciones farmacéuticas anteriores pueden usarse para prevenir el desarrollo de una enfermedad o trastorno (por ejemplo, una infección viral, una infección bacteriana, o cáncer) o para tratar un paciente aquejado con la enfermedad o trastorno (por ejemplo, una infección viral, una infección bacteriana, o cáncer). La enfermedad o trastorno puede diagnosticarse usando criterios generalmente aceptados en la materia. Por ejemplo, la infección viral puede diagnosticarse por la medición del título viral en una muestra del paciente, la infección bacteriana puede diagnosticarse detectando las bacterias en una muestra del paciente, y el cáncer puede diagnosticarse detectando la presencia de un tumor maligno. Las composiciones farmacéuticas pueden administrarse tanto antes de como tras la eliminación quirúrgica de tumores primarios y/o tratamiento tal como administración de radioterapia o fármacos quimioterapéuticos convencionales.

La inmunoterapia normalmente es inmunoterapia activa, en la que el tratamiento se basa en la estimulación *in vivo* del sistema inmunitario del huésped endógeno para reaccionar contra, por ejemplo, tumores o células bacterianamente o viralmente infectadas, con la administración de agentes modificadores de la respuesta inmunitaria (composiciones que comprenden ácidos nucleicos que codifican polipéptidos inmunogénicos como se proporciona en el presente documento).

La frecuencia de administración de las composiciones profilácticas o terapéuticas como se describe en el presente documento, además de la dosificación, variará de individuo a individuo, y puede establecerse fácilmente usando técnicas convencionales. Frecuentemente pueden administrarse entre 1 y 10 dosis durante un periodo de 52 semanas. Normalmente se administran 3 dosis, a intervalos de 1 mes, más normalmente, se administran 2-3 dosis cada 2-3 meses. Es posible que los intervalos sean más probablemente una vez al año para ciertas terapias. Las vacunaciones de refuerzo pueden administrarse periódicamente a partir de aquí. Protocolos alternos pueden ser apropiados para pacientes individuales y enfermedades y trastornos particulares. Una dosis adecuada es una cantidad de un compuesto que, cuando se administra como se ha descrito anteriormente, es capaz de promover, por ejemplo, una respuesta inmunitaria antitumoral, antiviral, o una antibacteriana, y está al menos el 10-50 % por encima del nivel basal (es decir, sin tratar). Tal respuesta puede monitorizarse midiendo los anticuerpos antitumorales en un paciente o por generación dependiente de vacuna de linfocitos T citolíticos capaces de destruir, por ejemplo, las células tumorales del paciente, las células viralmente infectadas del paciente, o las células bacterianamente infectadas del paciente *in vitro*. Tales vacunas también deben ser capaces de causar una respuesta inmunitaria que conduzca a un desenlace clínico mejorado (por ejemplo, remisiones más frecuentes, completas o parciales, o supervivencia libre de enfermedad más larga) en pacientes vacunados en comparación con pacientes no vacunados. Normalmente, la cantidad de los títulos virales estará entre $1,0 \times 10^4$ ufp/animal y $1,0 \times 10^{15}$ ufp/animal. Tamaños de dosis adecuada variarán con el tamaño del paciente, pero normalmente oscilarán de aproximadamente 0,01 ml a aproximadamente 10 ml, más normalmente de aproximadamente 0,025 a aproximadamente 7,5 ml, lo más normalmente de aproximadamente 0,05 a aproximadamente 5 ml. Aquellos expertos en la materia apreciarán que el tamaño de dosis puede ajustarse basándose en el paciente particular o la enfermedad o trastorno particular que está tratándose. Para administración por vía oral, el vector adenoviral quimérico puede formularse convenientemente en una píldora.

En general, una dosificación apropiada y pauta de tratamiento proporciona el (los) compuesto(s) activo(s) en una cantidad suficiente para proporcionar beneficio terapéutico y/o profiláctico. Una respuesta tal puede monitorizarse estableciendo un desenlace mejorado {por ejemplo, remisiones más frecuentes, supervivencia libre de enfermedad completa o parcial, o más larga) en pacientes tratados en comparación con pacientes no tratados. Tales respuestas inmunitarias pueden evaluarse generalmente usando ensayos de proliferación, citotoxicidad o citocinas estándar descritos anteriormente, que pueden realizarse usando muestras obtenidas de un paciente antes y después del tratamiento.

Por ejemplo, la detección de inmunocomplejos formados entre polipéptidos inmunogénicos y anticuerpos en líquido corporal que son específicos para polipéptidos inmunogénicos puede usarse para monitorizar la eficacia de terapia, que implica un polipéptido inmunogénico particular, para una enfermedad o trastorno en la que el polipéptido inmunogénico está asociado. Muestras de líquido corporal tomadas de un individuo antes de y posterior al inicio de la terapia pueden analizarse para los inmunocomplejos por las metodologías descritas anteriormente. Brevemente, se compara el número de inmunocomplejos detectados en ambas muestras. Un cambio sustancial en el número de inmunocomplejos en la segunda muestra (inicio post-terapia) con respecto a la primera muestra (pre-terapia) refleja

terapia satisfactoria.

A. Administración de las composiciones de la presente invención

5 Según la composición para su uso en métodos de la presente invención, una composición que comprende el vector
adenoviral quimérico se administra por vía oral, intranasalmente, o mucosamente, mediante, por ejemplo, la vagina,
pulmones, glándulas salivales, fosas nasales, intestino delgado, colon, recto, amígdalas o parches de Peyer. La
10 composición puede administrarse sola o con un adyuvante como se ha descrito anteriormente. En algunas
divulgaciones, los adyuvantes están codificados por una secuencia de ácidos nucleicos (por ejemplo, un ácido
nucleico que codifica IL-2, GM-CSF, IL-12, de flagelina bacteriana). En algunas divulgaciones, el adyuvante se
administra al mismo tiempo que la composición. En otras divulgaciones, el adyuvante se administra después de la
composición, por ejemplo, 6, 12, 18, 24, 36, 48, 60 o 72 horas después de la administración de la composición.

B. Detección de una respuesta inmunitaria a antígenos de interés

15 Una respuesta inmunitaria al polipéptido heterólogo puede detectarse usando cualquier medio conocido en la
materia que incluye, por ejemplo, detectar activación específica de linfocitos T CD4⁺ o CD8⁺ o detectar la presencia
de anticuerpos que se unen específicamente al polipéptido.

20 La activación específica de linfocitos T CD4⁺ o CD8⁺ asociados a una respuesta inmunitaria mucosa, humoral o
celular puede detectarse en una variedad de formas. Métodos de detección de la activación específica de linfocitos T
incluyen, pero no se limitan a, detectar la proliferación de linfocitos T, la producción de citocinas (por ejemplo,
linfocinas), o la generación de actividad citolítica (es decir, generación de linfocitos T citotóxicos específicos para el
25 polipéptido inmunogénico). Para linfocitos T CD4⁺, un método preferido de detección de la activación específica de
linfocitos T es la detección de la proliferación de linfocitos T. Para linfocitos T CD8⁺, un método preferido de
detección de la activación específica de linfocitos T es la detección de la generación de actividad citolítica usando
ensayos de liberación de ⁵¹Cr (véase, por ejemplo, Brossart y Bevan, Blood 90(4): 1594-1599 (1997) y Lenz et al., J.
Exp. Med. 192(8):1135-1142 (2000)).

30 La detección de la proliferación de linfocitos T puede llevarse a cabo mediante una variedad de técnicas conocidas.
Por ejemplo, la proliferación de linfocitos T puede detectarse midiendo la tasa de síntesis de ADN. Los linfocitos T
que se han estimulado para proliferar presentan una elevada tasa de síntesis de ADN. Una forma típica de medir la
tasa de síntesis de ADN es, por ejemplo, por cultivos de linfocitos T de marcado por pulsos con timidina tritiada, un
precursor de nucleósidos que se incorpora en ADN recientemente sintetizado. La cantidad de timidina tritiada
35 incorporada puede determinarse usando un espectrofotómetro de centelleo líquido. Otras formas de detección de la
proliferación de linfocitos T incluyen medir aumentos en la producción de interleucina-2 (IL-2), flujo de Ca²⁺
o captación de colorante, tal como 3-(4,5-dimetiltiazol-2-il)-2,5-difeniltetrazolio. Alternativamente, puede medirse la
síntesis de linfocinas (por ejemplo, interferón-gamma) o puede cuantificarse el número relativo de linfocitos T que
pueden responder al polipéptido inmunogénico.

40 Las respuestas inmunitarias de anticuerpos (también conocidas como respuestas inmunitarias humorales o
respuestas de linfocitos B), que incluyen respuestas de anticuerpos de la mucosa, pueden detectarse usando
inmunoensayos conocidos en la técnica [Tucker et al., Mol Therapy, 8, 392-399 (2003); Tucker et al., Vaccine, 22,
2500-2504 (2004)]. Inmunoensayos adecuados incluyen la técnica de inmunoensayo de sándwich de anticuerpo
45 monoclonal doble de David et al. (patente de EE.UU. N.º 4.376.110); ensayos de sándwich de anticuerpo
monoclonal-policlonal (Wide et al., en Kirkham y Hunter, eds., Radioimmunoassay Methods, E. y S. Livingstone,
Edinburgh (1970)); el método de "transferencia Western" de Gordon et al. (patente de EE.UU. N.º 4.452.901);
inmunoprecipitación de ligando marcado (Brown et al. (1980) J. Biol. Chem. 255:4980-4983); enzimoimmunoanálisis
de adsorción (ELISA) como se describen, por ejemplo, por Raines et al. (1982) J. Biol. Chem. 257:5154-5160;
50 técnicas inmunocitoquímicas, que incluyen el uso de fluorocromos (Brooks et al. (1980) Clin. Exp. Immunol. 39:477);
y neutralización de actividad (Bowen-Pope et al. (1984) Proc. Natl. Acad. Sci. USA 81:2396-2400). Además de los
inmunoensayos descritos anteriormente, están disponibles varios otros inmunoensayos, que incluyen aquellos
descritos en las patentes de EE.UU. N.º 3.817.827; 3.850.752; 3.901.654; 3.935.074; 3.984.533; 3.996.345;
4.034.074; y 4.098.876.

Ejemplos

Los siguientes ejemplos pretenden ilustrar, pero no limitar, la presente invención.

60 Ejemplo 1: Construcción de un vector adenoviral quimérico (DS1)

Para demostrar que los agonistas de TLR-3 pueden mejorar las respuestas inmunitarias adaptativas a antígenos
expresados de interés, se construyeron varios vectores adenovirales quiméricos diferentes que comprendían
secuencias de ácidos nucleicos que codificaban varios antígenos de interés diferentes. En este ejemplo, el ácido
65 nucleico que codifica gp120 (del NIH AIDS Reagent and Reference Reagent Program) se puso bajo el control de un
promotor del CMV con un pequeño intrón justo en la dirección 5' del codón de iniciación en el vector lanzadera

(pShuttle, Qbiogene). Se puso una cola de poli A de bGH en la dirección 3' del ácido nucleico que codifica gp120. La secuencia de vector se expone en SEQ ID NO: 1. Se realizó recombinación homóloga con el vector pAd (Qbiogene) para generar un vector capaz de producir Ad recombinante (delecionado E1/E3) que contenía el ácido nucleico que codifica gp120. Se generó DS1 transfectando la nueva construcción de expresión de pAd-CMV-gp120 en células 293. Los títulos se midieron por métodos convencionales.

Ejemplo 2: DS1 (vector más agonista de TLR-3) es superior a rAd5 estándar para inducir una respuesta inmunitaria específica de antígeno

Para determinar si la adición de agonista de TLR-3 podría mejorar las respuestas inmunitarias adaptativas, se administraron 10×10^7 UFP de cualquiera de rAd-CMV-gp120 más 5 ug/ml de poli I:C (DS1) o rAd-CMV-gp120 solo (rAd5) a animales por sonda nasogástrica oral en las semanas 0 y 3. Ambos vectores expresan gp120 del VIH bajo el control de un promotor del CMV y usan adenovirus delecionados en E1/E3 recombinantes tipo 5. Los títulos de anticuerpos para gp120 se midieron en el plasma 3 y 6 semanas después de la administración inicial por ELISA de de IgG anti-gp120 como se describe en Tucker, et al., *Mol Ther* 8:392 (2004)). Como se muestra en la Figura 1, DS1 cumplió significativamente mejor que rAd5 en provocar respuestas de anticuerpos a la proteína gp120 tanto a las 3 como a las 6 semanas después de la administración inicial por vía oral. En particular, el título de anticuerpos promedio para gp120 fue 100 veces mejor con el grupo de DS1 que con el grupo de rAd5 en la semana 6. También parece que el grupo de DS1 se reforzó por la readministración en la semana 4, porque el título promedio aumentó más de 20 veces entre las semanas 3 y 6 mientras que el grupo de rAd5 solo mostró un ligero aumento en el título medio de anticuerpos. Los resultados demuestran que la adición de un agonista de TLR-3 puede mejorar enormemente respuestas de anticuerpos mediadas por Ad5 a antígenos de interés tras la administración por vía oral de un vector adenoviral quimérico que comprende un ácido nucleico que codifica el antígeno de interés. Como control positivo para el ensayo también se midieron sueros de un animal inyectado subcutáneamente con gp120 más adyuvante completo de Freund en el ELISA de anti-gp120 en la semana 3. Los animales sin tratar y los animales administrados con el análogo de ARNbc solo (ARNbc) sirvieron de controles negativos y de fondo, respectivamente, para el ELISA. Cada grupo contuvo 6 animales.

Ejemplo 3: Construcción de un segundo vector adenoviral quimérico (DS1b) y un tercer vector adenoviral quimérico (DS1c)

Se insertó un ácido nucleico que codifica la proteína verde fluorescente (GFP) en pShuttle-CMV (Qbiogene) usando digestos de enzima de restricción estándar. El plásmido pShuttleCMV-GFP se combinó por recombinación homóloga con el vector pAd (Qbiogene) como se describe antes con el fin de generar un vector capaz de producir Ad recombinante (delecionado en E1/E3) que comprende una secuencia de ácidos nucleicos que codifica GFP. Se clonó un ácido nucleico que codifica hemaglutinina (HA) de la gripe A/PR8/34 y se puso en el vector pShuttle-CMV (Qbiogene) (SEQ ID NO: 13). El plásmido pShuttleCMV-HA (PR/8) se combinó por recombinación homóloga con el vector pAd (Qbiogene) como se describe antes con el fin de generar un vector capaz de producir Ad recombinante (delecionado en E1/E3) que comprende una secuencia de ácidos nucleicos que codifica HA. El Ad recombinante se generó transfectando la nueva construcción de expresión pAd-CMV-GFP y pAd-CMV-HA en células 293. Los títulos se midieron por métodos convencionales.

Ejemplo 4: DS1b (Ad-CMV-GFP más agonista de TLR-3) y DS1c (Ad-CMV-HA más agonista de TLR-3) es superior a rAd5 estándar para inducir una respuesta inmunitaria específica de antígeno.

Se administraron $1,0 \times 10^7$ UFP de cualquiera de Ad-CMV-GFP más 5 ug/ml de poli I:C (DS1b) o Ad-CMV-GFP (rAd5) a animales por sonda nasogástrica oral en la semana 0. Ambos virus expresan la GFP bajo el control de un promotor del CMV y usan adenovirus delecionados en E1/E3 recombinantes tipo 5. Los títulos de anticuerpos para GFP se midieron en el plasma 3 semanas después de la administración de virus inicial por ELISA de IgG anti-GFP. Como se muestra en la Figura 2, el grupo DS1b rindió significativamente mejor que rAd5 en provocar respuestas de anticuerpos a la proteína GFP 3 semanas después de la administración inicial por vía oral.

Las respuestas de linfocitos T CD8⁺ a GFP se midieron por tinción con tetrámeros de esplenocitos. Los animales se vacunaron en las semanas 0, 4, 8 y los bazos se recogieron en la semana 10. Los esplenocitos se tiñeron con CD8-FITC y el tetrámero que reconoce el epítipo inmunodominante para GFP en ratones Balb/c. Los resultados muestran que la administración por vía oral del vector DS1b fue estadísticamente mejor que rAd solo en inducir células CD8 positivas con tetrámeros (Figura 2b).

Se administraron $1,0 \times 10^7$ UFP de cualquiera de Ad-CMV-HA más 5 ug/ml de poli I:C (DS1c) o Ad-CMV-HA (rAd5) a animales por sonda nasogástrica oral en la semana 0. Ambos virus expresan HA bajo el control de un promotor del CMV y usan adenovirus delecionados en E1/E3 recombinantes tipo 5. Los títulos de anticuerpos para HA se midieron en el plasma 3 semanas después de la administración de virus inicial por ELISA de IgG anti-PR8/34. El procedimiento para medir las respuestas de anticuerpos es similar al descrito antes, con la excepción de que las placas de ELISA se recubrieron con 5 ug/ml de lisado de A/PR8/34 completo (Advanced Biotechnology Incorporated, Gaithersburg, MD). Como se muestra en la Figura 2C, el grupo DS1c rindió significativamente mejor (aproximadamente 100 veces mejor) que rAd5 en provocar respuestas de anticuerpos a la gripe 3 semanas después

de la administración inicial por vía oral. Los resultados de estos estudios también demuestran que el enfoque de uso de agonista de TLR-3 junto con un vector adenoviral recombinante quimérico puede aplicarse generalmente a múltiples antígenos heterólogos diferentes, con una mejora de 100 veces en el título de anticuerpos.

5 Ejemplo 5. Las vías de administración no parenterales son superiores a las vías parenterales

Se probó la administración intramuscular inyectando directamente $1,0 \times 10^7$ ufp de pAd-CMV-gp120 (DS10 +/- poli I:C) a 5 ug/ml en el cuádriceps de animales. Se midieron títulos de IgG en suero plasmático para GFP como se describe antes. Cada grupo contuvo 6 animales. Como se muestra en la Figura 3A, se observaron títulos de anticuerpos significativo para gp120 3 semanas después de la administración en el grupo con agonista de TLR-3 (rAd+PI i.m.). (Figura 3a).

Se probó la administración intranasal administrando 20 ul de $1,1 \times 10^6$ ufp de DS1c +/- 5 ug/ml de poli I:C en la fosa nasal de ratones. Los ratones se anestesiaron ligeramente con isoflurano antes de administrar el virus formulado en solución salina estéril. Los resultados muestran que rAd-CMV-HA más poli I:C (DS1c) tuvo títulos de anticuerpos ligeramente mayores en comparación con animales administrados con rAd-CMV-HA estándar. Los resultados se representan como animales individuales para los grupos de DS1c (N=6) y rAd (N=5). Se usan animales sin tratar (N=4) para controles negativos.

20 Ejemplo 6. Construcción de un agonista de TLR-3 expresado.

Se sintetizó un segmento de ADN corto de 45 pb ordenando oligos de ADN que cuando se hibridaron juntos formaron un segmento de 45 pb designado para preparar una horquilla de ARN bicatenario (**GAAACGATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTC**) (SEQ ID NO: 10). Este segmento corto (llamado luc1) se clonó en el plásmido pSK-que contiene el promotor de beta-actina humana y una cola de poli A de BGH. Este plásmido se llama pSK-luc1.

30 Ejemplo 7. Las funciones de pSK-luc 1 en cultivos celulares dendríticos como poli I:C, los efectos de poli I:C y rAd son aditivos.

Para determinar si el agonista de TLR-3 expresado del Ejemplo 6 anterior podría servir de inductor de citocinas pro-inflamatorias y maduración de células dendríticas como el ligando de TLR-3 poli I:C, se probó un agonista de TLR-3 de ARNbc expresado en cultivos celulares dendríticos. Se cultivaron médula ósea de los fémures de ratones Balb/c con ligando flt-3 (200 ng/ml), 5 % de suero, en medio DMEM, con el fin de preparar cultivos celulares dendríticos primarios. Cinco días después de establecerse los cultivos de médula ósea primarios, se transfectaron células 293 con tanto pSK-luc1, pSK-beta2 (un segmento largo de beta-galactosidasa que forma una horquilla de 200 pb) como pcDNA3 (vector de expresión vacío). En el día 6, las células transfectadas se trataron por irradiación UV (20 segundos a 40 kJ/cm^2) para producir la apoptosis y estas células se administraron a las células dendríticas. Cualquiera de poli I:C (1 ug/ml), rAd (1 ufp/célula), rAd + poli I:C, células transfectadas con pSK-luc1, células transfectadas con pSK-beta2, o células transfectadas con pcDNA3, se administraron a las células dendríticas y se cultivaron durante la noche. Como se muestra en la Figura 4A, las células transfectadas con pSK-luc1 pueden mejorar significativamente la activación de células dendríticas como se mide por el ELISA de IL-6 de ratón. Los resultados de este experimento también muestran que la combinación de rAd más ligando de TLR3 (poli I:C) juntos pueden mejorar enormemente la actividad de células dendríticas.

También se probaron ligandos adicionales. El agonista de TLR-3 expuesto en SEQ ID NO: 11 (m1) también forma una horquilla de ARNbc de aproximadamente el mismo tamaño que luc1. Esto se hizo solapando oligonucleótidos e hibridándolos juntos antes de la clonación en el vector pSK bajo el control del promotor de beta-actina humana. Los vectores se transfectaron en células 293 y se administran a células dendríticas primarias como se describe antes. Como se muestra en la Figura 4B, estos ligandos adicionales pueden activar células dendríticas similar a las del ligando luc1 (Figura 4B).

50 Ejemplo 8. Construcción de un cuarto vector adenoviral quimérico (DS2) y vectores de clonación rápida (DS2beta-luc y DS2C-luc).

Un ácido nucleico que codifica gp120 (del NIH AIDS Research and Reference Reagent Program) se puso bajo el control de un promotor del CMV con un intrón pequeño justo en la dirección 5' del codón de iniciación en el vector lanzadera (pShuttleCMV, Qbiogene). Se puso una cola de poli A de bGH en la dirección 3' del ácido nucleico que codifica gp120. El agonista de TLR-3 de ARNbc luc1 bajo el control del promotor de beta-actina humana y poli A (descritos en el Ejemplo 5 anteriormente) se insertó en el vector de gp120 pShuttle de forma que tanto el ácido nucleico que codifica gp120 como el ácido nucleico que codifica el agonista de TLR-3 estuvieran contenidos en un único vector bajo el control de dos promotores separados. La orientación de la expresión del ácido nucleico que codifica el antígeno de interés y la expresión del agonista de TLR-3 se ilustra en la Figura 5.

65 También se construyeron dos vectores lanzadera genéricos llamados DS2beta-luc (SEQ ID NO: 14) y DS2C-luc (SEQ ID NO: 15) de forma que un ácido nucleico que codifica cualquier antígeno de interés pudiera insertarse bajo

un promotor del CMV y cualquiera del promotor de beta-actina humana o un promotor del CMV se usara para conducir la expresión de un agonista de TLR-3 de ARNbc. En particular, el vector DS2C-luc tiene un sitio Kpn 1 único en el que puede clonarse fácilmente un ácido nucleico que codifica un antígeno de interés. El fin de estos vectores es hacer la posterior construcción de vectores mucho más fácil debido a que un ácido nucleico que codifica cualquier antígeno de interés podría insertarse en el sitio de clonación para fabricar rápidamente un vector capaz de provocar respuestas de anticuerpos y linfocitos T contra el antígeno de interés. La recombinación homóloga de DS2 con el pAd de vector (Qbiogene) se realizó como antes con el fin de generar un vector capaz de producir Ad recombinante (delecionado en E1/E3) que contenía un ácido nucleico que codifica GFP y un ácido nucleico que codifica el agonista de TLR-3 de ARNbc luc1. Se generó Ad recombinante transfectando la nueva construcción de expresión pAd-beta-actina-luc1-CMV-gp120 en células 293. Los títulos se midieron por métodos convencionales.

Ejemplo 9: Inducción de una respuesta inmunitaria específica de antígeno tras la administración oral de DS2.

Se administraron $1,0 \times 10^7$ UFP de cualquiera de pAd-CMV-gp120 más el agonista de TLR-3 luc1 (DS2) o pAd-CMV-gp120 (rAd5) a animales por sonda nasogástrica oral en la semana. Ambos virus expresan el gp120 bajo el control de un promotor del CMV y usan adenovirus delecionados en E1/E3 recombinantes tipo 5. Los títulos de anticuerpos para gp120 se midieron en el plasma 3 semanas después de la administración de virus por ELISA de IgG anti-gp120. El protocolo de ELISA se ha descrito antes (Tucker, et al, Mol Therapy 8:392 (2004)). Los resultados demuestran que DS2 puede inducir aproximadamente una mejora de 2 log en el título de anticuerpos para gp120, el antígeno heterólogo usado en el experimento. El vector de DS2 comprende una secuencia de ácidos nucleicos que codifica que expresa gp120 y una secuencia de ácidos nucleicos que expresa un agonista de TLR-3 de ARNbc. Como control positivo para el ensayo también se midieron sueros de dos animales inyectados subcutáneamente con 10 microgramos de proteína gp120 más adyuvante completo de Freund en el ELISA de anti-gp120. Los animales sin tratar sirvieron de controles negativos para el ELISA. Cada grupo contuvo 6 animales. Los resultados se ilustran en la Figura 6.

Ejemplo 10: Inducción de una respuesta inmunitaria específica de antígeno tras la administración oral de DS3.

Se administraron $1,0 \times 10^7$ UFP de cualquiera de pAd-CMV-HA de la gripe (de A/PR/8/34) más el ligando de TLR7/8 ácido poliuridílico (DS3) o pAd-CMV-HA (rAd5) a animales por sonda nasogástrica oral en la semana 0. Ambos virus expresan HA de la gripe bajo el control de un promotor del CMV y usan adenovirus delecionados en E1/E3 recombinantes tipo 5. Los títulos de anticuerpos para HA se midieron en el plasma 3 semanas después de la administración de virus por ELISA de IgG anti-HA de la gripe. Cada grupo contuvo 6 animales. Los resultados se ilustran en la Figura 7.

Ejemplo 11: Construcción de un quinto, un sexto y un séptimo vector adenoviral quimérico (DS2b, DS2b-for y ND1.1 214).

Se sintetizó el gen HA de la gripe (A/Indo/5/2005) por CelTek (Nashville, TN) y se dispuso en el vector pShuttle-CMV (Qbiogene) que tiene un promotor del CMV con un intrón pequeño justo en la dirección 5' del codón de iniciación en el vector lanzadera. El ADN de luc1 con promotor de beta-actina humana y poli A (descrito en el Ejemplo 5) se colocaron en el vector en la dirección 3' del antígeno, en la orientación mostrada en la Figura 5 para DS2b. La secuencia de luc1 es (**GAAACGATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTC**) (SEQ ID NO: 10) y el vector pShuttle completo se expone en SEQ ID NO: 6. Una orientación alternativa de luc1 con el promotor en un vector lanzadera se describe como SEQ ID NO: 7 y se designa DS2b-for. Los presentes inventores también han construido otro vector pShuttle (llamado DS2bC-HA) (SEQ ID NO: 16) que comprende dos promotores del CMV separados que conducen la expresión del agonista de TLR-3 luc1 y HA de la gripe descritos anteriormente. La recombinación homóloga con el vector pAd (Qbiogene) se realizó como antes con el fin de generar vectores capaces de producir Ad recombinante (delecionados en E1/E3) que contenía el ácido nucleico que codifica HA y el agonista de TLR-3 luc1 bajo promotores separados. Se generó Ad recombinante transfectando las nuevas construcciones de pAd en células 293. Los títulos se midieron por métodos convencionales. El vector de pAd completado que contiene DS2C-luc se llamó ND1.1 214 y se depositó en el depósito de patentes de ATCC el 22 de febrero de 2007 (Manassus, VA). La secuencia de ácidos nucleicos de este vector adenoviral quimérico se expone en SEQ ID NO: 17. El ácido nucleico que codifica el antígeno heterólogo está en negrita y está flanqueado por un sitio de reconocimiento Cla I sobre el extremo 5' y un sitio de reconocimiento Not 1 sobre el extremo 3'. La secuencia de ácidos nucleicos que codifica los agonistas de TLR-3 está en cursiva, con la secuencia de conector en negrita. Una secuencia de ácidos nucleicos que codifica cualquier antígeno de interés y una secuencia de ácidos nucleicos que codifica cualquier agonista de TLR-3 expresado adecuado puede insertarse en el vector adenoviral quimérico.

Ejemplo 12: Inducción de una respuesta inmunitaria específica de antígeno tras la administración oral de DS2b.

Se administraron $1,0 \times 10^7$ UFP de cualquiera de pAd-CMV-HA más el agonista de TLR-3s luc1 en la orientación inversa (DS2b) u orientación directa (DS2b-for), o pAd-CMV-HA (rAd5), a animales por sonda nasogástrica oral en la semana 0. Estos virus expresan el antígeno HA de la gripe bajo el control de un promotor del CMV y usan adenovirus delecionados en E1/E3 recombinantes tipo 5. Los títulos de anticuerpos para HA se midieron en el plasma 3 semanas después de la administración de virus por ELISA de IgG anti-HA. Los resultados demuestran que

5 el vector de DS2b provoca una respuesta de anticuerpos a la proteína HA superior al vector de rAd estándar (rAd5). El vector de DS2b contiene rAd5 que expresa HA, además de expresar un agonista de receptor 3 tipo Toll (TLR3), una horquilla de ARN bicatenario, demostrando que el uso del ligando de ARNbc codificado puede mejorar respuestas inmunitarias adaptativas a antígenos de interés. Como se muestra en la Figura 8A y la Figura 6, el ARNbc expresado puede mejorar respuestas inmunitarias adaptativas a múltiples antígenos heterólogos diferentes. Los animales sin tratar sirvieron de control negativo para el ELISA. Cada grupo contuvo 6 animales.

10 Se examinaron vectores en la orientación opuesta (DS2for) para respuestas de anticuerpos tras tanto administración oral como intramuscular de $1,0 \times 10^7$ ufp de virus por animal a 0 y 5 semanas. Las respuestas de anticuerpos a HA se midieron 4 y 7 semanas después de la administración inicial. Como se muestra en la Figura 8B, el vector de orientación opuesta también pueden inducir respuestas de anticuerpos sustanciales a antígenos heterólogos. DS1b y DS1b para los vectores indujeron respuestas similares a HA en el momento de tiempo de 4 semanas. Significativamente, el efecto de refuerzo de la respuesta de anticuerpos se demostró con DS1b para el vector y mostró que podrían usarse múltiples dosis para aumentar las respuestas de anticuerpos al antígeno heterólogo.

15 También se demostró otro ejemplo de potencial del enfoque de vectores adenovirales quiméricos. El vector ND1.1 214 se administró a animales por administración oral ($1,0 \times 10^7$ ufp) o intranasal (3×10^6 ufp) y las respuestas de anticuerpos al antígeno heterólogo se midieron en la semana 3. Como se muestra en la Figura 8C, se midieron respuestas de anticuerpos sustanciales a HA tras la administración por vía oral, muy más allá de los valores típicos de una administración única por vía oral de vector de rAd.

20

Listado de secuencias informal

SEQ ID NO: 1 (DS1)

taacatcatc aataatatac cttatTTTTg attgaagcca atatgataat
 gagggggtg agtttgtgac gtggcgcggg gcgtgggaac ggggCGGGTg
 acgtagtagt gtggcggaag tgtgatgTtg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgTtt ttggTgtgCG cCGTgtaca
 caggaagtga caatTTtcgc gcggtTTtag gcggatgTtg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtgaaa
 tctgaataatTTTgtgttactcatagcgcgtaataactggtaccgCGGCGcctcgagtctag
 agatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTCCC
 AGTCACGACGTTGTA AAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTG
 GGTACTGGCCAcagagcTtgcccattgcatacgtTgtatccatatacataatgtacattt
 atattggctcatgtccaacattaccgccaTgttgacattgattattgactagttattaatag
 taatcaattacggggTcattagTtcatagccatataTggagTtccgcgttacataacttac
 ggtaaatggcccCGcctggctgaccgcccacgacccccgcccattgacgtcaataatgacgt
 atgttcccatagtaacgccaatagggactTtccattgacgtcaatgggtggagTatttacgg
 taaactgcccactTggcagTacaTcaagtgtatcatatgccaagTaccgcccattgacgt
 caatgacggTaaatggcccCGcctggcattatgcccagTacaTgacctatgggactTtccTa
 ctTggcagTacaTctacgtattagTcatcgctattaccatggTgatgcggTttTggcagTac
 atcaatgggCGTggatagcggTttgactcacggggattTccaagTctccaccccattgacgt
 caatgggagTttgtttTggcaccaaaatcaacgggactTtccaaaatgtcgtaacaactcCG
 cccattgacgcaaatgggCGTtaggcgtgTaccggtgggaggtctatataagcagagctcgt
 ttagTgaaccgTcagatcgctggagacgccatccacgctgtttTgacctccatagaagaca
 cgggaccgatccagcctgactctagCctAGCTCtgaagtggTggTgaggccctgggCagg
 ttggTatcaaggTtacaagacaggTttaaggagaccaatagaaactgggcatgtggagacag
 agaagactctTgggtttctgataggcactgactctctctgcctattggTctatTTtcccacc
 cttaggctgctggTctgagcctagGAGATCTCTCGAGGTCGACGGTATAGCTTCTAGAGATC
 CCTCGACCTCGAGATCCATTGTGCTCTAAAGGAGATACCCGGCCAGACACCCTCACCTGCGG
 TGCCCAGCTGCCAGGCTGAGGCAAGAGAAGGCCAGAAACCATGCCATGGGGTCTCTGCAA
 CCGCTGGCCACCTTGTACCTGCTGGGGATGCTGGTCGCTTCCGTGCTAGCTGTGGAGAAGCT
 GTGGGTGACTGTATACTATGGGGTGCCTGTGTGGAAGGAGGCCACCACCACCCTGTTCTGTG
 CCTCTGATGCCAAGGCCATGACACTGAGGTCCACAATGTCTGGGCCACCCATGCCTGTGTG
 CCCACTGACCCCAACCCTCAGGAGGTGGTGTGGAGAATGTGACTGAGCACTTCAACATGTG
 GAAGAACAACATGGTGGAGCAGATGCAGGAGGACATCATCAGCCTGTGGGACCAGAGCCTGA
 AGCCCTGTGTGAAGCTGACCCCCCTGTGTGTGACCCTGAACTGCAAGGATGTGAATGCCACC
 AACACCACCAATGACTCTGAGGGCACTATGGAGAGGGGTGAGATCAAGAACTGCAGCTTCAA
 CATCACCACCAGCATCAGGGATGAGGTGCAGAAGGAGTATGCCCTGTTCTACAAGCTGGATG
 TGGTGCCCATTGACAACAACAACACCAGCTACAGGCTGATCAGCTGTGACACCTCTGTGATC
 ACCCAGGCCTGCCCAAGATCAGCTTTGAGCCATCCCCATCCACTACTGTGCCCTGCTGG
 CTTTGCCATCCTGAAGTGCAATGACAAGACCTTCAATGGCAAAGGCCCTTGAAGAATGTGA
 GCACTGTGCAGTGCACCTCATGGCATCAGGCCTGTGGTGTGAGCACCAGCTGCTGCTGAATGGC
 AGCCTGGCTGAGGAGGAGGTGGTGTGATCAGGTCTGACAACCTTACCAACAATGCCAAGACCAT
 CATTGTGCAGC
 TGAAGGAGTCTGTGGAGATCAACTGCACCAGGCCAACACAACACCAGGAAGAGCATTAC
 ATTGGCCCTGGCAGGGCCTTCTACACCCTGGGGAGATCATTGGGGACATCAGGCAGG
 CCCACTGCAACATCAGCAGGGCCAAGTGAATGACACCCTGAAGCAGATTGTGATCAAGCTG
 AGGGAGCAGTTTGAAGAACAAGACCATTGTGTTCAATCACAGCTCTGGTGGTGTGATCCTGAGAT
 TGTGATGCACAGCTTCAACTGTGGTGGTGTGAGTTCTTCTACTGCAACAGCACCAGCTGTTCA
 ACAGCACCTGGAACAACAACACTGAGGGCAGCAACAACACTGAGGGCAACACCATCACCTG

CCTTGCAGGATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTATGCTCC
TCCCATCAGGGGCCAGATCAGGTGCAGCAGCAACATCACTGGCCTGCTGCTGACCAGGGATG
GTGGCATCAATGAGAATGGCACTGAGATTTTCAGGCCTGGTGGTGGGGACATGAGGGACAAC
TGGAGGTCTGAGCTGTACAAGTACAAGGTGGTGAAGATTGAGCCCTTGGTGTGGCTCCAC
CAAGGCTAAGCGCAGGGTGGTGCAGAGGGAGAAGcgcgctgtgggctgaggatccCGAGGGT
GAGTGTCTGCTGGACGCATCCCGGCTATGCAGCCCCAGTCCAGGGCAGCAAGGCAGGCC
CCGTCTGCCTCTTACCCEGGAGCCTCTGCCCGCCCCACTCATGCTCAGGGAGAGGGTCTTCT
GGCTTTTCCCAGGCTCTGGGCAGGCACAGGCTAGGTGCCCTAACCCAGGCCCTGCACACA
AAGGGGCAGGTGCTGGGCTCAGACCTGCCAAGAGCCATATCCGGGAGGACCCTGCCCTGAC
CTAAGCCCACCCCAAAGGCCAAACTCTCCACTCCCTCAGCTCGGACACCTTCTCTCCTCCCA
GATTCAGTAACTCCCAATCTTCTCTCTGCAGAGCCCAAATCTTGTGACAAAACCTCACACAT
GCCACCGTGCCCAGGTAAGCCAGGCCAGGCCCTCGCCCTCCAGCTCAAGGCGGGACAGGTGC
CCTAGAGTAGCCTGCATCCAGGGACAGGCCCCAGCCGGGTGCTGACACGTCACCTCCATCT
CTTCTCAGCACCTGAACCTCTGGGGGGACCGTCAGTCTTCTCTTCCCCCAAACCCAAG
GACACCTCATGATCTCCCGGACCCCTGAGGTACATGCGTGGTGGTGGACGTGAGCCACGA
AGACCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAA
AGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAAGCGTCTCACCGTCTGCAC
CAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGCCCC
CATCGAGAAAACCATCTCCAAGCCAAAGGTGGGACCCGTGGGGTGGCAGGGGCCACATGGAC
AGAGGCCGGCTCGGCCACCCTCTGCCCTGAGAGTGACCGCTGTACCAACCTCTGTCCTACA
GGGCAGCCCCGAGAACCACAGGTGTACACCTGCCCCATCCCGGGATGAGCTGACCAAGAA
CCAGGTGAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGG
AGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGTGGACTCCGACGGC
TCCTTCTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTT
CTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGT
CTCCGGGTAAATGAGTGCAGCGGCCCGCAGGTAAAGCCAGCCCAGGCCCTCGCCCTCCAGCTCAA
GGCAGGACAGGTGCCCTAGAGTAGCCTGCATCCAGGGACAGGCCCCAGCCGGGTGCTGACAC
GTCCACCTCCATCTCTTCTCAGGTCTGCCCGGGTGGCATCCCTGTGACCCCTCCCAAGTGC
CTCTCTGGCCCTGGAAAGTTGCCACTCCAGTGCCACCAAGCCTTGTCTAATAAAATTAAGT
TGCATCATTTTGTCTGACTAGGTGTCTTCTATAATATTATGGGGTGGAGGGGGTGGTATG
GAGCAAGGGGCCAAAGTTAACTTGTATTATGCAGCTTATAATGGTTACAAATAAAGCAATAG
CATCACAAATTTACAAATAAAGCATTTTTCCTACTGCATTCTAGTTGTGGTTTGTCCAAAC
TCATCAATGTATCTTATCATGTCTGgatctggg cgtgggtaag

421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
tgtatctgtt ttgcagcgc

481 cgccgcccgc atgagcacca actcgtttga tggagcatt
gtgagctcat atttgacaac

541 ggcgatgccc ccattgggccc ggggtgcgtca gaatgtgatg
ggctccagca ttgatggctg

601 cccgctcctg cccgcaaac ctactacctt gacctacgag
accgtgtctg gaacgccgtt

661 ggagactgca gcctccgccc ccgcttcagc cgctgcagcc
accgcccgcg ggattgtgac

721 tgactttgct ttctgagcc cgcttgcaag cagtgcagct
tcccgttcat ccgcccgcga

781 tgacaagtgg acggetcttt tggcacaatt ggattctttg
accgggaac ttaatgtcgt

841 ttctcagcag ctgttggatc tgcgccagca ggtttctgcc
ctgaaggctt cctcccctcc

901 caatgcccgtt taaaacataa ataaaaaac agactctgtt
tggatttggg tcaagcaagt

961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcggctctg
 1021 gtcgttgagg gtccctgtgta tttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccgt ctctggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cgggggtggg ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggg gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttggact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagotta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tcactaacgt catagttgtg
 1561 ttccaggatg agatcgatc atggccatttt tacaagcgc
 gggcggaggg tgccagactg
 1621 cgggtataatg gttccatccg gccccagggg gtagttacc
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgccc
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct gggagaaaag caggttccctg
 agcagctgcg acttaccgca
 1801 gccgggtggg ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggccac ttcgttaagc
 atgtccctga ctgcatgtt
 1921 ttccctgacc aaatccgcca gaaggcgtc gccgcccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacgggt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcggctcc acagctcggg cacctgctct
 acggcatctc gatccagcat
 2101 atctctctgt ttcgcccgtt ggggcggcct tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacgggcca gggtcatgtc tttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acgggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctctg
 2281 ctgggtgctga agcgtgccc gtcttcgccc tgcgctcgg
 ccaggtagca tttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgcccacg aggggcagtg cagacttttg agggcgtaga
 gcttgggccc gagaaatacc

2461 gattccgggg agtaggcac cgcgccgcag gccccgcaga
 cggctctcgca ttccacgagc
 2521 caggtgagct ctggccggtc ggggtcaaaa accaggtttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagttta aacgaattca
 atagcttgtt gcatgggcgg
 2701 cgatataaaa tgcaagggtc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac cttataagc
 2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaata
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcgggtc gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgtc
 3241 cagaacaaca tacagcgtt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca ccagaaaac cgcacgcgaa cctacgcca
 gaaacgaaag caaaaaaacc
 3481 cacaacttcc tcaaatcgtc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgcctaa
 aacctagtc acccgccccg
 3601 ttcccacgcc ccgcgccacg tcacaaactc caccctca
 ttatcatatt ggcttcaatc
 3661 caaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg cttcctcgt
 3781 cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaag
 3901 ccagcaaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg

3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcggg gtaggtcggt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggat gtaggcgggt ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtggttttt ttgtttcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaacg aaaactcacg ttaagggatt
 ttggatcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttgggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gccccagtgc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtggctc
 4921 tgcaacttta tccgcctcca tccagtctat taattggtgc
 cgggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccgatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt ttcgcatgat tgaacaagat
 ggattgcagc caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc

5461 tgatgccgcc gtgttccggc tgtcagcgca ggggcgcccg
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggcggt ccttgccgag ctgtgctcga cgttgteact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgcccagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatcga gcgagcacgt
 actcggatgg aagccggctt
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gcgatgcctg
 5941 cttgccgaat atcatgggtg aaaatggccg cttttctgga
 ttcategact gtggccggct
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttcctcgt gctttacggt
 atcgccgctc ccgattcgca
 6121 ggcgatcgc ttctatcgc ttcttgacga gttcttctga
 attttgtaa aatttttgtt
 6181 aatcagctc attttttaac caataggccg aatcggcaa
 catcccttat aatcaaaag
 6241 aatagaccgc gataggggtg agtgttgttc cagtttgaa
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccaactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgagggtgccg
 taaagtcta aatcggaacc
 6421 ctaaaggag cccccgattt agagcttgac ggggaaagcc
 ggcgaacgtg gcgagaaagg
 6481 aaggaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgctcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID: 2 (Secuencia del vector DS2)

taacatcacc aataatatac cttattttgg attgaagcca atatgataat
 gagggggtgg agtttgtgac gtggcgcggg gcgtgggaac gggcggggtg
 acgtagtagt gtggcggaag tgtgatgttg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgttt ttgggtgtgc cgggtgtaca
 caggaagtga caattttcgc gcggttttag gcggatgttg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg
 taatact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**

CAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
 TAGGGCGAATTGGGTACTGGCCAcagagccttggccattgcatacgttggatccatatacata
 atatgtacatttatattggctcatgtccaacattaccgcatggtgacattgattattgact
 agttattaatagtaatacaattacggggcattagttcatagcccatatattggagttccgcgt
 tacataacttacggtaaatggcccgcctggctgaccgccaacgacccccgccattgacgt
 caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtatttacggtaaaactgcccaactggcagtacatcaagtgtatcatatgccaaagtacgcc
 ccctattgacgtcaatgacggtaaatggcccgcctggcattatgccagtacatgacctat
 gggactttcctacttggcagtacatctacgtattagtcacogctattaccaatgggtgatgagg
 ttttggcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctcca
 cccattgacgtcaatgggagtttggtttggcaccaaaatcaacgggactttccaaaatgtc
 gtaacaactccgccccattgacgcaaattgggcggttaggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgctcagatcgccctggagacgccatccacgctggttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagtgggtggtag
 gccctgggcaggtgggtatcaaggttacaagacaggtttaaggagaccaatagaact**gggc**
atgtggagacagagaagactccttgggtttctgataggcactgactctctctgcctattggc
 tattttcccacccttaggctgctggctgagcctag**GAGATCTCTCGAGGTCGACGGTATCG**
ATGGGTACAGCTTCTAGAGATCCCTCGACCTCGAGATCCATTGTGCTCTAAAGGAGATACCC
 GGCCAGACACCCTCACCTGCGGTGCCAGCTGCCAGGCTGAGGCAAGAGAAGGCCAGA AAC
 CATGCCATGGGGTCTCTGCAACCGCTGGCCACCTTGTACCTGCTGGGGATGCTGGTCTGCTT
 CCGTGCTAGCTGTGGAGAAGCTGTGGGTGACTGTATACTATGGGGTGCCTGTGTGGAAGGAG
 GCCACCACCACCCTGTTCTGTGCCTCTGATGCCAAGGCCTATGACACTGAGGTCCACAATGT
 CTGGGCCACCATGCCTGTGTGCCACTGACCCCAACCCTCAGGAGGTGGTCTGGAGAATG
 TGACTGAGCACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCAGGAGGACATCATC
 AGCCTGTGGGACCAGAGCCTGAAGCCCTGTGTGAAGCTGACCCCTGTGTGTGACCCTGAA
 CTGCAAGGATGTGAATGCCACCAACACCACCAATGACTCTGAGGGCACTATGGAGAGGGGTG
 AGATCAAGAACTGCAGCTTCAACATCACCACCAGCATCAGGGATGAGGTGCAGAAGGAGTAT
 GCCCTGTTCTACAAGCTGGATGTGGTGCCATTGACAACAACAACACCAGCTACAGGCTGAT
 CAGCTGTGACACCTCTGTGATCACCCAGGCCTGCCCAAGATCAGCTTTGAGCCCATCCCCA
 TCCACTACTGTGCCCTGCTGGCTTTGCCATCCTGAAGTGCAATGACAAGACCTTCAATGGC
 AAAGGCCCTTGCAAGAATGTGAGCACTGTGCAGTGCACCTCATGGCATCAGGCCTGTGGTGAG
 CACCCAGCTGCTGCTGAATGGCAGCCTGGCTGAGGAGGAGGTGGTGTGATCAGGTCTGACAAC
 TCACCAACAATGCCAAGACCATCATTGTGCAGC
 TGAAGGAGTCTGTGGAGATCAACTGCACCAGGCCCAACAACAACACCAGGAAGAGCATTAC
 ATTGGCCCTGGCAGGGCCTTCTACACCCTGGGGAGATCATTGGGGACATCAGGCAGG
 CCCACTGCAACATCAGCAGGGCCAAAGTGAATGACACCCTGAAGCAGATTGTGATCAAGCTG
 AGGGAGCAGTTTGAAGAACAAGACCATTGTGTTCAATCACAGCTCTGGTGGTGTGATCCTGAGAT
 TGTGATGCACAGCTTCAACTGTGGTGGTGAAGTCTTCTACTGCAACAGCACCCAGCTGTTCA
 ACAGCACCTGGAACAACAACACTGAGGGCAGCAACAACACTGAGGGCAACACCATCACCCCTG
 CCTTGACAGGATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTATGCTCC
 TCCCATCAGGGGCCAGATCAGGTGCAGCAGCAACATCACTGGCCTGCTGCTGACCAGGGATG
 GTGGCATCAATGAGAATGGCACTGAGATTTTCAGGCCTGGTGGTGGGGACATGAGGGACAAC
 TGGAGGTCTGAGCTGTACAAGTACAAGGTGGTGAAGATTGAGCCCTTGGTGTGGCTCCCAC
 CAAGGCTAAGCGCAGGGTGGTGCAGAGGGAGAAGcgcgctgtgggctgaggatccCGAGGGT
 GAGTGCTCCTGCCTGGACGCATCCCGGCTATGCAGCCCCAGTCCAGGGCAGCAAGGCAGGCC
 CCGTCTGCCTCTTACCCGGAGCCTCTGCCCGCCCCACTCATGCTCAGGGAGAGGGTCTTCT
 GGCTTTTTCCAGGCTCTGGGCAGGCACAGGCTAGGTGCCCTAACCAGGCCCTGCACACA
 AAGGGCAGGTGCTGGGCTCAGACCTGCCAAGAGCCATATCCGGGAGGACCCTGCCCTGAC
 CTAAGCCACCCCAAGGCCAAACTCTCCACTCCCTCAGCTCGGACACCTTCTCTCCTCCCA
 GATTCCAGTAACCTCCCAATCTTCTCTCTGCAGAGCCCAATCTTGTGACAAAACCTCACACAT
 GCCACCGTGCCAGGTAAGCCAGCCAGGCCTCGCCCTCCAGCTCAAGGCGGGACAGGTGC

CCTAGAGTAGCCTGCATCCAGGGACAGGCCCCAGCCGGGTGCTGACACGTCCACCTCCATCT
CTTCCTCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCCTTTCCCCCAAACCCAAG
GACACCCTCATGATCTCCCGGACCCCTGAGGTACATGCGTGGTGGTGGACGTGAGCCACGA
AGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAA
AGCCCGGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTACAGCGTCCACACCGTCTGCAC
CAGGACTGGCTGAATGGCAAGGAGTACAAGTGAAGGTCTCCAACAAAGCCCTCCCAGCCCC
CATCGAGAAAACCATCTCCAAGCCAAAGGTGGGACCCGTGGGGTGGCAGGGCCACATGGAC
AGAGGCCGGCTCGGCCACCCTCTGCCCTGAGAGTGACCGCTGTACCAACCTCTGTCTTACA
GGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCATCCCGGGATGAGCTGACCAAGAA
CCAGGTGAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGG
AGAGCAATGGGCAGCCGGAGAACAATAAGACCACGCCTCCCGTGTGGACTCCGACGGC
TCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAAAGTCTT
CTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGT
CTCCGGGTAAATGAGTGGCAGCGCCGCAGGTAAGCCAGCCCAGGCCCTCGCCCTCCAGCTCAA
GGCGGGACAGGTGCCCTAGAGTAGCCTGCATCCAGGGACAGGCCCCAGCCGGGTGCTGACAC
GTCCACCTCCATCTCTTCTCAGGTCTGCCCGGGTGGCATCCCTGTGACCCCTCCCCAGTGC
CTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCACCAGCCTTGTCTAATAAAATTAAGT
TGCATCATTGTTGCTGACTAGGTGTCCTTCTATAATATTATGGGGTGGAGGGGGGTGGTATG
GAGCAAGGGGCCAAGTTAACTTGTATTTATGAGCTTATAATGGTTACAAATAAAGCAATAG
CATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAAC
TCATCAATGTATCTTATCATGTCTGgatctggg cgtgggtaag

421 ggtgggaaag aatatataag gtgggggtct tatgtagttt tgtatctggt
ttgcagcagc

481 cgccgcccgc atgagcacca actcgtttga tggaagcatt gtgagctcat

CGCGGGCCGCCCTATTCTATAGTGTACCTAAATGCTA**GAGCTC**GCTGATCAGCCTCGACTG
TGCCTT
CTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
ACTCCCAC
TGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
TGGGGGGT
GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
GGTGGGCT
CTATGGCTTCTGAGGCGGAAAGAACCTATGGCTTCTGAGGCGGAAAGAACA**CA**ACCACCGCGG
TGGCGGCCGCCACACAAAAACCAACACACAGATGTAATGAAATAAAGATATTTTATTTCT
AGAGAAACGATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTCTGCAGGAATT
CGCCCTTTAGATATCATCGATGTCTCGGCGGTGGTGGCGCGTCGCGCCGCTGGGTTTTATAG
GGCGCCGCCGCGGCCGCTCGAGCCATAAAAGGCCA**ACT**TTCGGAACGGCGCACGCTGATTGGC
CCCAGCGCCGCTCACTCACCGGCTTCGCCGCACAGTGCAGCATTTTTTTACCCCTCTCCCT
CCTTTTGCGAAAAAAAAGAGCGAGAGCGAGATTGAGGAAGAGGAGGAGGGAGAGTTTTG
GCGTTGGCCGCCTTGGGGTGTGGGCGTCGACGATATCTAAGGGCGAATTCGATATCAagct
agct**Tgtcgactcg** aagatctggg cgtgggtaag

421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
tgtatctggt ttgcagcagc

481 cgccgcccgc atgagcacca actcgtttga tggaagcatt
gtgagctcat attgacaac

541 gcgcatgccc ccatgggccc ggggtgctca gaatgtgatg
ggctccagca ttgatggtcg

601 ccccgctcctg cccgcaaact ctactacctt gacctacgag
 accgtgtctg gaacgccggt
 661 ggagactgca gcctccgccc cgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac
 721 tgactttgct ttctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga
 781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggaac ttaatgtcgt
 841 ttctcagcag ctggtggatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcggtt taaaacataa ataaaaaacc agactctggt
 tggatttggga tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcggctctg
 1021 gtcggtgagg gtctgtgta tttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccgt ctctggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cggggtgggt ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttgact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcgccct gggcgaagat atttctggga
 tcaactaacgt catagtgtg
 1561 ttccaggatg agatcgtcat aggccatttt taciaagcgc
 gggcggaggg tgccagactg
 1621 cggataaatg gttccatccg gcccaggggc gtagttacc
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct ggaagaagag caggttctctg
 agcagctgcg acttaccgca
 1801 gccggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aatccgccga gaaggcgctc gccgccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacgggt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcggtccc acagctcggg cacctgctct
 acggcatctc gatccagcat

2101 atctcctcgt ttcgcgggtt ggggcggctt tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacgggcca gggatcatgtc tttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctctg
 2281 ctgggtgctga agcgcctgcc gtcttcgccc tgcgcgtcgg
 ccaggtagca tttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc cttggaggag
 2401 ggcggcagc aggggcagtg cagacttttg agggcgtaga
 gcttggggcg gagaaatacc
 2461 gattccgggg agtaggcac cgcggcagc gccccgcaga
 cggctctcga ttccacgagc
 2521 caggtgagct ctggccgttc ggggtcaaaa accaggttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagtta aacgaattca
 atagcttgtt gcatgggcg
 2701 cgatataaaa tgcaagggtc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac cttataagc
 2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaata
 3001 agcaccacg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcgggtc gtgctaaaaa ggcaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgctc
 3241 cagaacaaca tacagcgtt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca ccagaaaaac cgcacgcgaa cctacgcccc
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaatcgtc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgcctaa
 aacctagtc acccgccccg

3601 ttcccacgcc ccgcgccaag tcacaaactc cccccctca
 ttatcatatt ggcttcaatc
 3661 caaaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg cttcctcgct
 3781 cactgactcg ctgcgctcgg tcggtcggct gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcggttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctccctteg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcggg gtaggtcggt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggat gtaggcggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttgga tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtgggtttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaacg aaaactcacg ttaagggatt
 ttggatcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttggctctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gccccagtgc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtggctc
 4921 tgcaacttta tccgctcca tccagtctat taattgttgc
 cgggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc

ES 2 573 456 T3

5101 ccggatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggocggtt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt ttcgcatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgcc gtgttccggc tgtcagcgca ggggccccg
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggctt ccttgccgag ctgtgctcga cgttgtcact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccg ggcaggatct cctgtcatct
 caccttgctc ctgccgagaa
 5701 agtatccatc atggctgatg caatgccggc gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatcga gcgagcacgt
 actcggatgg aagccggtct
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gogatgcctg
 5941 cttgccgaat atcatggtgg aaaatggccg cttttctgga
 ttcacgcact gtggccggct
 6001 ggggtgtggc gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttcctcgt gctttacggt
 atcgccgctc ccgattcgca
 6121 gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga
 attttgttaa aatttttgtt
 6181 aaatcagctc attttttaac caataggccg aaatcggcaa
 catcccttat aaatcaaaag
 6241 aatagaccgc gataggggtg agtgttgttc cagtttgaa
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgaggtgccg
 taaagctcta aatcgggaacc
 6421 ctaaagggag cccccgattt agagcttgac ggggaaagcc
 ggcgaacgtg gcgagaaagg
 6481 aagggaaaga agcgaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgcgtcca ttcgccatc
 6601 aggatcgaat taattcttaa t

SEQ ID No: 3 GAAACGATATGGGCTGAATACTTAAGTATTCAGCCCATATCGTTTC.

5 Seq id no: 4

CGGGCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCGCCCTTAGATATC
GTCGACGCCAGCACCCCAAGGCGGCCAACGCCAAACTCTCCCTCCTCCTCTCCTCAATC
TCGCTCTCGCTCTTTTTTTTTTTCGCAAAGGAGGGGGAGAGGGGGTAAAAAATGCTGCACT
GTGCGGCGAAGCCGGTGAGTGAGCGGCGGGGCCAATCAGCGTGCGCCGTTCCGAAAGTTG
CCTTTTATGGCTCGAGCGGCCGCGGGCGCCCTATAAAACCCAGCGGCGGACGCGCCACC
ACCGCCGAGACATCGATGATATCTA
AAGGGCGAATTCCTGCAGCCCGGGGATCCACTAGTCTAGATGCATGCTCGAGCGGCCGCCA
GTGTGATGGATATCTGCAGAATTCGCCCTCAGCTGCGGATCCATTCGCCATTCAGGCTGCG
CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG
GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTA
ACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCT
CGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGGATCCACTAGTT
TCTAGAAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGGCGGCCG
CCACCGCGGTGGAGCT
Atcgaattcaagcttgtcgactc gaagatc
CT AG ACTAGTGGAT CCCCCGGGCTGCAGGAATT
CGCCCTTTAGATATCATCGATGTCTCGGCGGTGGTGGCGCGTCCGCGCCGCTGGGTTTTATAG
GGCGCCCGCGCGGCCGCTCGAGCCATAAAAGGCAACTTTCGGAACGGCGCACGCTGATTGGC
CCCGCGCCGCTCACTCACCGGCTTCGCCGCACAGTGCAGCATTTTTTTACCCCTCTCCCT
CCTTTTGCGAAAAAAAAGAGCGAGAGCGAGATTGAGGAAGAGGAGGAGGGAGAGTTTTG
GCGTTGGCCGCTTGGGGTGCTGGGCGTCGACGATATCTAAGGGCGAATT C
GATATCAAGC TTATCGATAC CGTCGACCTC GAGGGGGGC CCG

SEQ NO: 5

CGGGCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATC
GAATTCGCCCTTAGATATCGTCGACGCCAGCACCCCAAGGCGGCCA
ACGCCAAACTCTCCCTCCTCCTCTCCTCAATCTCGCTCTCGCTCTTTTTTTTTTTCGCAA
AAGGAGGGGAGAGGGGGTAAAAAATGCTGCACTGTGCGGCGAAGCCGGTGAGTGAGCGGCG
CGGGGCCAATCAGCGTGCGCCGTTCCGAAAGTTGCCTTTTATGGCTCGAGCGGCCGCGGGCG
CGCCCTATAAAACCCAGCGGCGGACGCGCCACCACCGCCGAGACATCGATGATATCTA
AAGGGCGAATTCCTGCAGCCCGGGGATCCACTAGTCTAGATGCATGCTCGAGCGGCCGCCA
GTGTGATGGATATCTGCAGAATTCGCCCTT
CAGCTGCGGATCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGC
GATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGT
GCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGACG
TTGTAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAAT
TGGGTACCGGGCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATC
GAATTCCTGCAGCCCGGGGATCCACTAGTT
TCTAGAAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGGCGGCCG
CCACCGCGGTGGAGCTA
Tcgaattcaagcttgtcgactc gaagatcgtaca
caggaagtgacaatcttcgcgcggttttaggcgatggtg
tagtaaatttggcgtaaccgagtaagatttggccatcttcgcgggaaaactgaataaga

ggaagtgaaatctgaataatTTTTgtgttactcatagocgtaataactggtac
 CGGGCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATC
 GAATTCGCCCTTAGATATCGTCGACGCCAGCACCCCAAGGCGGCCA
 ACGCCAAACTCTCCCTCCTCCTCCTCAATCTCGCTCTCGCTCTTTTTTTTTTCGCAA
 AAGGAGGGGAGAGGGGGTAAAAAATGCTGCACTGTGCGGCGAAGCCGGTGAGTGAGCGGCG
 CGGGGCCAATCAGCGTGCGCCGTTCCGAAAGTTGCCTTTTATGGCTCGAGCGGCCGCGGCGG
 CGCCCTATAAAACCCAGCGGCGCGACGCGCCACCACCGCCGAGACATCGATGATATCTA
 AAGGGCGAATTCTGCAGCCCGGGGGATCCACTAGTCTAGAACTAGTGGATCCCCCGGGCTG
 CAGGAATTC
 GATATCAAGCTTATCGATAACCGTCGACCTCGAGGGGGGGCCCGGTACCCA
 ATTCGCCCTATAGTGAGTCGTATTACAATCACTGGCCGTGTTTTACAA
 CGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTAATCGCCTTGACAGC
 ACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGGATC
 GCCCTTCCAACAGTTGCGCAGCCTGAATGGCGAATGGATCCGCAGCTG
 AAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCA
 TCTAGAAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGGCGGCCG
 CCACCGCGGTGGAGCTA

SEQ ID: 6

taacatcatc aataatatac cttatTTTTgg attgaagcca atatgataat
 gagggggtgg agTTTTgtgac gtggcgcggg gcgtgggaac ggggcgggtg
 acgtagtagt gtggcggaaag tgtgatggtg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgttt ttggtgtgog ccggtgtaca
 caggaagtga caatTTTcgc gcggtTTtag gcggatggtg tagtaaattt
 gggcgtaacc gagtaagatt tggccatTTT cgcgggaaaa ctgaataaga
 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg
 taataact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**
 CAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
 TAGGGCGAATTGGGTACTGGCCAcagagcttggcccattgcatacgttgtatccatatac
 atatgtacatTTtatattggctcatgtccaacattaccgccaatggtgacattgattattgact
 agttattaatagtaatcaattaacggggtcattagttcatagccatataatggagttccgcgt
 tacataacttacggtaaatggcccgcctggctgaccgccaacgacccccgcccattgacgt
 caataatgacgtatggtcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtatttacggtaaaactgcccacttggcagtacatcaagtgtatcatatgccaaagtacgcc
 cccattgacgtcaatgacggtaaatggcccgcctggcattatgccagtacatgacctat
 gggactttcctaacttggcagtacatctac**gtaattagtcacogctattacca**tggtgatgagg
 ttttggcagtacatcaatgggcgtggatagcgglttgactcacggggatttccaagtctcca
 ccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtc
 gtaacaactccgccccattgacgcaaatgggcggtagggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgtcagatcgctggagacgccatccacgctgttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttggtgggtgag
 gccctgggcaggttggatcaaggttacaagacaggtttaaggagaccaatagaaact**gggc**
atgtggagacagagaagactccttgggtttctgataggcaactgactctctctgacctattggtc
 tattttcccacccttaggctgctggtctgag**cctagGAGATCTCTCGAGGTCGACGGTAT**cg
 atgcca ccATGGAGAA AATCGTCTTG TGTCTGCTA TTGTGTCTCT

- 51 AGTGAAGAGC GATCAAATTT GTATCGGCTA CCATGCCAAT AACTCAACAG
- 101 AGCAGGTCGA TACTATCATG GAGAAAAACG TAACAGTTAC TCATGCCCAA
- 151 GACATCTTGG AAAAGACCCA CAACGGCAAA CTTTGGCACC TGGATGGAGT
- 201 GAATCCCCTG ATCCTCCGGG ACTGTTCACT CGCTGTTGG CTGCTCGGGA
- 251 ACCCTATGTG TGATGAGTTT ATCAACGTGC CTGAATGGTC TTACATTGTG
- 301 GAGAAGGCTA ACCCTACCAA TGACCTCTGC TATCCTGGGT CATTAAACGA

351 TTACGAGGAA CTGAAACACC TGTGTCTAG AATTAACCAC TTTGAAAAGA
 401 TACAGATTAT ACCCAAGTCT AGTTGGAGTG ATCACGAAGC CTCCTCAGGC
 451 GTTAGCTCAG CGTGTCCCTA TCTGGGCTCT CCATCCTTCT TTAGAAATGT
 501 GGTCTGGTTA ATCAAAAAGA ACAGTACCTA CCCAACCATC AAAAAGTCTT
 551 ATAACAATAC CAATCAGGAG GACCTGCTCG TGTGTGGGG TATCCATCAC
 601 CCGAACGACG CCGCTGAACA GACTAGGCTG TATCAGAACC CCACTACATA
 651 CATCAGTATT GGCACGAGTA CTCTGAACCA GCGATTAGTG CCAAAGATTG
 701 CAACACGGAG CAAAGTAAAT GGGCAATCTG GCAGGATGGA GTTTTTCTGG
 751 ACAATCTTAA AACCCAACGA TGCATAAAT TTCGAGTCCA ATGGCAATTT
 801 CATCGCCCCT GAATACGCCT ATAAGATCGT GAAAAAGGGG GACTCTGCAA
 851 TTATGAAGTC CGAATTAGAG TATGGCAATT GCAACACGAA GTGCCAGACA
 901 CCAATGGGAG CCATTAATAG CTCAATGCCC TTCCATAATA TTCATCCATT
 951 GACCATTGGG GAGTGCCCAA AGTACGTGAA GTCCAACCGC CTGGTCTCTG
 1001 CAACCGGTCT AAGAAATAGC CCGCAGAGAG AATCGCGGAG GAAGAAACGT
 1051 GGCCTGTTTG GCGCGATTGC CGGATTCATC GAGGGAGGCT GGCAGGGTAT
 1101 GGTCGATGGT TGGTACGGAT ACCACCATAG CAACGAACAG GGGTCCGGCT
 1151 ATGCAGCAGA TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAC
 1201 AAGGTTAATA GTATTATAGA TAAATGAAC ACGCAATTCG AGGCCGTTGG
 1251 GAGGGAGTTT AACAACTGG AACGCCGGAT CGAAAATCTG AATAAGAAAA
 1301 TGGAAGACGG CTTCTTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA
 1351 CTCATGGAGA ACGAGAGGAC CCTGGATTTC CACGATAGCA ACGTGAAGAA
 1401 CCTTTACGAC AAGGTGAGAC TTCAGCTCCG AGACAACGCC AAGGAGCTGG
 1451 GGAATGGATG CTTCGAGTTT TACCACAAAT GTGACAATGA GTGCATGGAA
 1501 AGTATACGCA ACGGGACCTA CAATTACCCT CAGTATAGCG AAGAGGCTCG
 1551 GCTCAAACGC GAAGAGATAA GCGGGGTGAA ATTGGAATCA ATCGGAACAT
 1601 ATCAAATCCT GTCCATCTAT TCCACCGTCG CCTCTTCGCT GGCCCTCGCT
 1651 ATCATGATGG CTGGTCTGTC CCTATGGATG TGTTCCAATG GAAGCCTTCA
 1701 GTGCCGTATT TGTATATGAg c

GGCCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCC
 TT
 CTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCCAC
 TGTCTTTCCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
 TGGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCT
 CTATGGCTTCTGAGGGCGAAAGAAC**CAA**

CCACCGCGGTGGCGGCCGCCACACAAAAACCAACACACAGATGTAATGAAAATAAAGATAT
 TTTATTTCTAGAGAAACGATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTCCT
 GCAGGAATTCGCCCTTTAGATATCATCGATGTCTCGCGGTGGTGGCGCGTCGCGCCGCTGG
 GTTTTATAGGGCGCCGCGCGCCGCTCGAGCCATAAAAGGCAACTTTCGGAACGGCGCACG
 CTGATTGGCCCCGCGCCGCTCACTCACCGGCTTCGCCGCACAGTGCAGCATTTTTTTACCCC
 CTCTCCCCTCCTTTTTCGAAAAAAAAGAGCGAGAGCGAGATTGAGGAAGAGGAGGGG
 AGAGTTTTGGCGTTGGCCGCCTTGGGGTGTGGGCGTCGACGATATCTAAGGGCGAATTCGA
 TATCA

AGCTAGCT **Tgtcgactcg** aagatctggg cgtgggtaag
 421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctggt ttgcagcagc

481 cgccgccc atgagcacca actcgtttga tggaaacatt
 gtgagctcat atttgacaac
 541 gcgcatgccc ccatgggccc ggggtgcgtca gaatgtgatg
 ggctccagca ttgatggctg
 601 ccccgctctg cccgcaaact ctactacctt gacctacgag
 accgtgtctg gaacgccgtt
 661 ggagactgca gcctcccgcg ccgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac
 721 tgactttgct ttctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga
 781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggaac ttaatgtcgt
 841 ttctcagcag ctgttggatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcggtt taaaacataa ataaaaaac agactctggt
 tggatttggga tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcgtctcgc
 1021 gtcggtgagg gtcctgtgta tttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccgt ctctggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cggggtggtg ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttgact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tcaactaacgt catagttgtg
 1561 ttccaggatg agatcgtcat aggccathtt tacaagcgc
 gggcggaggg tgccagactg
 1621 cggataaatg gttccatccg gccaggggc gtagttaccc
 tcacagattt gcatttcca
 1681 cgctttgagt tcagatggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggttc
 1741 cgggtaggg gagatcagct ggaagaaag caggttcctg
 agcagctgcg acttaccgca
 1801 gccggtgggc ccgtaaata cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca gggggccac ttcgtaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aaatccgcca gaaggcgtc gccgcccagc
 gatagcagtt cttgcaagga

1981 agcaaagttt ttcaacgggt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcgggtccc acagctcggg cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttcgcggggt ggggcggctt tcgctgtacg
 gcagtagtgc gtgctcgtcc
 2161 agacgggcca gggtcattgtc tttccacggg cgcaggggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctgggtcctg
 2281 ctggtgctga agcgtgcgcg gtcttcgccc tgcgcgtcgg
 ccaggtagca tttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgcgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggccc gagaaatacc
 2461 gattccgggg agtaggcatt cgcgcgcag gccccgcaga
 cggctctgca ttccacgagc
 2521 caggtgagct ctggccgttc ggggtcaaaa accaggttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagtta aacgaattca
 atagcttggt gcatgggccc
 2701 cgatataaaa tgcaagggtc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac cttataagc
 2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaaaa
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcgggtca gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgtc
 3241 cagaacaaca tacagcgtt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca cccagaaaac cgcacgcgaa cctacgccc
 gaaacgaaag ccaaaaaacc

3481 cacaacttcc tcaaatcgtc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgccctaa
 aacctacgtc acccgccccg
 3601 ttcccacgcc ccgcgccacg tcacaaactc caccocctca
 ttatcatatt ggcttcaatc
 3661 caaataaagg tatattattg atgatgtaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg cttcctcgtc
 3781 cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagttag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctcctc
 gtgcgctctc ctgctccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctccctcgg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcggg ttaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggat gttaggcggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca acaaaaccac cgctggtagc
 ggtggttttt ttgtttgcga
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cttttgatct tttctacggg
 4561 gtctgacgct cagtggaacg aaaactcag ttaagggatt
 ttggatcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttgggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta ttctggttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gccccagtgc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtggctc
 4921 tgcaacttta tccgctcca tccagtctat taattggtgc
 cgggaagcta gagtaagtag

4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccggatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggocggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt ttcgcatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgcc gtgttccggc tgtcagcgca ggggccccg
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggagtt ccttgccgag ctgtgctcga cgttgtcact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgccgagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatcga gcgagcacgt
 actcggatgg aagccggtct
 5821 tgtegatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggetcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gcgatgcctg
 5941 cttgccgaat atcatggtgg aaaatggccg cttttctgga
 ttcategact gtggccggct
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttctcgt gctttacggt
 atcgccgctc ccgattcgca
 6121 gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga
 attttgttaa aatttttggt
 6181 aatcagctc attttttaac caataggccg aatcggcaa
 catcccttat aatcaaaag
 6241 aatagaccgc gatagggttg agtgttgttc cagtttgaa
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccaactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgagggtgccg
 taaagctcta aatcggaacc
 6421 ctaaagggag cccccgattt agagcttgac ggggaaagcc
 ggcgaacgtg gcgagaaagg

6481 aaggaagaa agcgaagga gcggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacacccgog cgcttaatgc gccgctacag
 ggcgctcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID: 7

taacatcatc aataatatac cttatfttgg attgaagcca atatgataat
 gagggggtgg agtttgtgac gtggcgcggg gcgtgggaac gggcggggtg
 acgtagtagt gtggcggaag tgtgatgttg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgttt ttggtgtgog ccggtgtaca
 caggaagtga caatfttcgc gcggtfttag gcggatgttg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg
 taatact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**
CAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
TAGGGCGAATTGGGTACTGGCCAcagagcttggcccattgcatacgttgtatccatata
 atatgtacatttatattggctcatgtccaacattaccgccatgttgacattgattattgact
 agttattaatagtaatcaattacgggggtcattagttcatagcccatatatggagttccgcgt
 tacataacttacggtaaatggcccgcctggctgaccgccaacgacccccgccattgacgt
 caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtatttacggtaaatggcccgcctggcagtacatcaagtgtatcatatgccaagtacgcc
 cctattgacgtcaatgacggtaaatggcccgcctggcattatgccagtacatgacctat
 gggactttcctacttggcagtacatctac**gattattagtcatogetattaccat**ggtgatgcgg
 ttttggcagtacatcaatggggtggatagcggtttgactcacggggatttccaagtctcca
 cccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgic
 gtaacaactccgccccattgacgcaaatgggcggtagggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgctcagatcgccctggagacgccatccacgctgtttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttggtggtgag
 gccctgggcaggttggatcaaggttacaagacaggtttaaggagaccaatagaaact**gggc**
atgtggagacagagaagactccttgggtttctgataggcactgactctctctgcctattggtc
 tattttcccacccttaggctgctggtctgag**ectagGAGATCTCTCGAGGTCGACGGTATcg**
 atgcca ccATGGAGAA AATCGTCCTG TTGCTCGCTA TTGTGTCTCT

51 AGTGAAGAGC GATCAAATTT GTATCGGCTA CCATGCCAAT AACTCAACAG
 101 AGCAGGTCGA TACTATCATG GAGAAAAACG TAACAGTTAC TCATGCCCAA
 151 GACATCTTGG AAAAGACCCA CAACGGCAAA CTTTGCGACC TGGATGGAGT
 201 GAAGCCCCTG ATCCTCCGGG ACTGTTTAGT CGCTGGTTGG CTGCTCGGGA
 251 ACCCTATGTG TGATGAGTTT ATCAACGTGC CTGAATGGTC TTACATTGTG
 301 GAGAAGGCTA ACCCTACCAA TGACCTCTGC TATCCTGGGT CATTTAACGA
 351 TTACGAGGAA CTGAAACACC TGTTGTCTAG AATTAACCAC TTTGAAAAGA
 401 TACAGATTAT ACCCAAGTCT AGTTGGAGTG ATCACGAAGC CTCCTCAGGC
 451 GTTAGCTCAG CGTGTCCCTA TCTGGGCTCT CCATCCTTCT TTAGAAATGT
 501 GGTCTGGTTA ATCAAAAAGA ACAGTACCTA CCCAACCATC AAAAAGTCTT
 551 ATAACAATAC CAATCAGGAG GACCTGCTCG TGTTGTGGGG TATCCATCAC
 601 CCGAACGACG CCGCTGAACA GACTAGGCTG TATCAGAACC CCACTACATA
 651 CATCAGTATT GGCACGAGTA CTCTGAACCA GCGATTAGTG CCAAAGATTG
 701 CAACACGGAG CAAAGTAAAT GGGCAATCTG GCAGGATGGA GTTTTTCTGG
 751 ACAATCTTAA AACCCAACGA TGCGATAAAT TTCGAGTCCA ATGGCAATTT
 801 CATCGCCCCT GAATACGCCT ATAAGATCGT GAAAAAGGGG GACTCTGCAA
 851 TTATGAAGTC CGAATTAGAG TATGGCAATT GCAACACGAA GTGCCAGACA

901 CCAATGGGAG CCATTAATAG CTCAATGCC TTCCATAATA TTCATCCATT
 951 GACCATTGGG GAGTGCCCAA AGTACGTGAA GTCCAACCGC CTGGTCCTCG
 1001 CAACCGGTCT AAGAAATAGC CCGCAGAGAG AATCGCGGAG GAAGAAACGT
 1051 GGCTGTTTG GCGGATTGC CGGATTCATC GAGGGAGGCT GGCAGGGTAT
 1101 GGTCGATGGT TGGTACGGAT ACCACCATAG CAACGAACAG GGGTCCGGCT
 1151 ATGCAGCAGA TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAAC
 1201 AAGGTTAATA GTATTATAGA TAAAATGAAC ACGCAATTCG AGGCCGTTGG
 1251 GAGGGAGTTT AACCAATCTGG AACGCCGGAT CGAAAATCTG AATAAGAAAA
 1301 TGAAGACGG CTTCCTTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA
 1351 CTCATGGAGA ACGAGAGGAC CCTGGATTTT CACGATAGCA ACGTGAAGAA
 1401 CCTTTACGAC AAGGTGAGAC TTCAGCTCCG AGACAACGCC AAGGAGCTGG
 1451 GGAATGGATG CTTCGAGTTT TACCACAAAT GTGACAATGA GTGCATGGAA
 1501 AGTATACGCA ACGGGACCTA CAATTACCCT CAGTATAGCG AAGAGGCTCG
 1551 GCTCAAACGC GAAGAGATAA GCGGGGTGAA ATTGGAATCA ATCGGAACAT
 1601 ATCAAATCCT GTCCATCTAT TCCACCGTCG CCTCTTCGCT GGCCCTCGCT
 1651 ATCATGATGG CTGGTCTGTC CCTATGGATG TGTCCAATG GAAGCCTTCA
 1701 GTGCCGTATT TGTATATGAg c

GGCCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCC
TT
 CTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCAC
 TGTCTTTCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
 TGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGCT
 CTATGGCTTCTGAGGCGGAAAGAACCAA

TGATATCGAATTCCGCCCTTAGATATCGTCGACGCCAGCACCCCAAGGCGGCCAACGCCAAA
 ACTCTCCCTCCTCCTCCTCCTCAATCTCGCTCTCGCTCTTTTTTTTTTTTCGAAAAGGAGGG
 GAGAGGGGTAAAAAATGCTGCACTGTGCGGCGAAGCCGGTGAGTGAGCGGCGCGGGGCCA
ATCAGCGTGCGCCGTTCCGAAAGTTGCCTTTTATGGCTCGAGCGGCCGCGGCCGCCCTAT
AAAACCCAGCGGCGCGACGCCACCACCGCCGAGACATCGATGATATCTAAAGGGCGAATT
 CCTGCAGgaaacgatatgggctgaatacggatccgtattcagccatatacgtttc**TCTAGAA**
 ATAAAATATCTTTATTTTATTACATCTGTGTGTGGTTTTTTGTGTGGCGGCCGCCACCGC
 GTGG

agctTgtcgaotcg aagatctggg cgtgggtaag
 421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctggt ttgcagcagc
 481 cgccgccc atgagacca actcgttga tgaagcatt
 gtgagctcat atttgacaac
 541 ggcgatgcc ccatgggccc ggggtgcgtca gaatgtgatg
 ggetccagca ttgatggtcg
 601 cccgctctg cccgcaaact ctactacctt gacotacgag
 accgtgtctg gaacgcccgtt
 661 ggagactgca gcctccgccc cgcttcagc cgetgcagcc
 accgcccgcg ggattgtgac
 721 tgactttgct ttctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga
 781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggaac ttaatgtcgt

841 ttctcagcag ctggttgatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcgggt taaaacataa ataaaaaacc agactctggt
 tggatttggga tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcggctctg
 1021 gtcgttgagg gtcctgtgta ttttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccgt ctctgggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cgggggtgggtg ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtgggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttggact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tcactaacgt catagtgtg
 1561 ttccaggatg agatcgatc atggccatttt taaaaagcgc
 gggcggaggg tgccagactg
 1621 cgggtataatg gttccatccg gcccaggggc gtagttacc
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct gggagaagaa caggttcctg
 agcagctgcg acttaccgca
 1801 gccgggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aatccgcca gaaggcgtc gccgcccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacggtt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcgtccc acagctcgggt cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttcgcgggtt ggggcggctt tcgctgtacg
 gcagtagtgc gtgctcgtcc
 2161 agacgggcca gggtcatgtc tttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acgggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctctg
 2281 ctgggtgctga agcgtgccc gtcttcgccc tgcgcgtcgg
 ccaggtagca tttgaccatg

2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggegc
 gcagcttgcc cttggaggag
 2401 gcgccgcacg aggggcagtg cagacttttg agggcgtaga
 gcttggggcg gagaaatacc
 2461 gattccgggg agtaggcac cgcgccgcag gccccgcaga
 cggctctcgca ttccacgagc
 2521 caggtgagct ctggccgctt ggggtcaaaa accaggtttc
 ccccatgctt tttgatgctt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagtta aacgaattca
 atagcttgtt gcatgggagg
 2701 cgatataaaa tgcaagggtc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac cttataagc
 2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaana
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcgggtc gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggg
 aaaatagcac cctcccgctc
 3241 cagaacaaca tacagcgtt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 aaaaaaaca cccagaaaac cgcacgcgaa cctacgcccc
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaatcgtc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgcctaa
 aacctacgtc acccgccccg
 3601 ttcccacgcc ccgcgccacg tcacaaactc caccacctca
 ttatcatatt ggcttcaatc
 3661 caaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcgggtgta
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg ctctctcgtt
 3781 cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg
 gtatcagctc actcaaaggc

3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggegt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcggg gtaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggtat gtaggcgggt ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtggttttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cttttgatct tttctacggg
 4561 gtctgacgct cagtggaacg aaaactcacg ttaagggatt
 ttggtcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttgggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gccccagtgc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtggctc
 4921 tgcaacttta tccgcctcca tccagtctat taattgttgc
 cgggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccggatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca

5341 agagacagga tgaggatcgt ttcgcatgat tgaacaagat
 ggattgcacg caggttctec
 5401 ggccgcttgg gtggagagge tattcggeta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgce gtgttccggc tgtcagecga ggggcgcccg
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggcggt ccttgccgag ctgtgctcga cgttgctact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgccgagaa
 5701 agtatccatc atggctgatg caatgccggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatcga gcgagcacgt
 actcggatgg aagccggtct
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gcgatgcctg
 5941 cttgccgaat atcatggtgg aaaatggccg cttttctgga
 ttcacgcact gtggccggct
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttcctcgt gctttacggt
 atcgcgcctc ccgattcgca
 6121 gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga
 attttgttaa aatttttggt
 6181 aatcagctc attttttaac caataggccg aatcggcaa
 catcccttat aatcaaaag
 6241 aatagaccgc gatagggttg agtgttgttc cagtttgaa
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccaactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgaggtgccg
 taaagctcta aatcggaacc
 6421 ctaaaggag cccccgattt agagcttgac ggggaaagcc
 ggcgaaactg gcgagaaagg
 6481 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgcgcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID:8 GATGGTGCTTCAAGCTAGTACttaaGTA CTAGCTTGAAGCACCATC
 SEQ ID: 9 GATGGTGCTTCAAGCTAGTACggatccGTA CTAGCTTGAAGCACCATC
 SEQ ID: 10 GAAACGATATGGGCTGAATACggatccGTATTCAGCCCATATCGTTTC
 SEQ ID: 11 CCTAATAATTAT CAAAATGTggatccACATTTTGATAATTATTAGG
 SEQ ID: 12 CCTAATAATTAT CAAAATGTaattACATTTTGATAATTATTAGG
 SEQ ID: 13 (DS1c)

1 taacatcatc aataatatac cttatTTTTGG attgaagcca atatgataat
 gagggggTGG
 61 agTTTTgtgac gtggcgcggg gcgtgggaac ggggcgggtg
 acgtagtagt gtggcggaag
 121 tgtgatgTTG caagtgtggc ggaacacatg taagcgacgg
 atgtggcaaaa agtgacgTTT
 181 ttggTgtgCG cggTgtaca caggaagtga caatTTTTcgC
 gcggTTTTtag gcggatgTTG
 241 tagtaaattt gggcgtaacc gagtaagatt tggccatttt
 cgcgggaaaa ctgaataaga
 301 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg
 taatactgta atagtaatca
 361 attacggggT cattagttca tagcccatat atggagttcc
 gcgttacata acttacggta
 421 aatggcccgC ctggctgacc gcccaacgac ccccgcccat
 tgacgtcaat aatgacgtat
 481 gttcccatag taacgccaat agggactttc cattgacgtc
 aatgggtgga gtattttacgg
 541 taaactgccc acttggcagt acatcaagtg tatcatatgc
 caagtacgcc ccctattgac
 601 gtcaatgacg gtaaattggcc cgccctggcat tatgcccagt
 acatgacctt atgggacttt
 661 cctacttggc agtacatcta **cgTattagtc atcgctatta**
ccatgggtgat gcggTTTTgg
 721 cagtacatca atgggcgtgg atagcggttt gactcacggg
 gatttocaag tctccacccc
 781 attgacgtca atgggagttt gTTTTggcac caaaatcaac
 gggactttcc aaaatgtcgt
 841 aacaactccg ccccatgac gcaaatgggc ggtaggcgTg
 tacggTggga ggtctatata
 901 agcagagctg gtttagtgaa cogtcagatc cgctagagat
ctGTTACCGAGCTCGGATCCGCCACCATGGAGGCAAACCTACTGGTCCTGTTATGTGCACTT
GCAGCTGCAGATGCAGACACAATATGTATAGGCTACCATGCGAACAAATTCACCGACACTGG
TGACACAGTActcgagaaaga atgtgacagt **gacacactct gttaacctgc**
 181 tcgaagacag ccacaacgga aaactatgta gattaaaagg
 aatagcccca ctacaattgg
 241 ggaaatgtaa catcgccgga tggctcttgg gaaaccaga
 atgcgaccca ctgcttccag
 301 tgagatcatg gtccctacatt gtagaaacac caaactctga
 gaatggaata tgttatccag
 361 gagatttcat cgactatgag gagctgaggg agcaattgag
 ctcagtgtca tcattcgaaa
 421 gattcgaaat atttcccaaa gaaagctcat ggcccaacca
 caacacaacc aaaggagtaa
 481 cggcagcatg ctcccatgCG gggaaaagca gTTTTtacag
 aaatttgcta tggctgacgg

541 agaaggaggg ctcatacca aagctgaaaa attccttatgt
 gaacaagaaa gggaaagaag
 601 tccttgact gtggggatt catcaccgt ctaacagtaa
 ggatcaacag aatatctatc
 661 agaatgaaaa tgcttatgtc tctgtagtga cttcaaatta
 taacaggaga tttaccccg
 721 aatagcaga aagacccaaa gtaagagatc aagctgggag
 gatgaactat tactggacct
 781 tgctaaaacc cggagacaca ataattttg aggcaaatgg
 aatctaata gcaccaaggt
 841 atgctttcgc actgagtaga ggctttgggt cggcatcat
 cacctcaaac gcatcaatgc
 901 atgagtgtaa cacgaagtgt caaacacccc tgggagctat
 aaacagcagt ctccctttcc
 961 agaataata cccagtcaca ataggagagt gcccaata
 cgtcaggagt gccaaattga .
 1021 ggatggttac aggactaagg aacattccgt ccattcaatc
 cagaggtcta tttggagcca
 1081 ttgccggttt tattgaagg ggatggactg gaatgataga
 tggatggtag ggttatcatc
 1141 atcagaatga acaggatca ggctatgcag cggatcaaaa
 aagcacacaa atgccatta
 1201 acgggattac aaacaaggtg aactctgtta tcgagaaaat
 gaacattcaa ttcacagctg
 1261 tgggtaaaga attcaacaaa ttagaaaaaa ggatggaaaa
 tttaaataaa aaagttgatg
 1321 atggatttct ggacatttgg acatataatg cagaattggt
 agttctactg gaaaatgaaa
 1381 ggactctgga tttccatgac tcaaatgtga agaattctgta
 tgagaaagta aaaagccaat
 1441 taaagaataa tgccaaagaa atcggaaatg gatgttttga
 gttctaccac aagtgtgaca
 1501 atgaatgcat ggaaagtgta agaaatggga cttatgatta
 tcccaaatat tcagaagagt
 1561 caaagttgaa cagggaaaag gtagatggag tgaaattgga
 atcaatgggg atctatcaga
 1621 ttctggcgat ctactcaact gtcgccagtt cactggtgct
 tttggtctcc ctgggggcaa
 1681 tcagtttctg gatgtgttct aatggatctt tgcagtgcag
 aatatgcac tgagattaga
 1741 atttcagaga tatgaggaaa aacacccttg tttctact
CCCAAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAA
CGCAGTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCAC
CCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCCGGCCTCTTGCGGGATGG
GCGGCCGCtcgagcct aagcttctag ataagatc cgatccaccg gatctagata
 actgatcata
 1021 atcagccata ccacatttgt agaggtttta cttgctttaa
 aaaacctccc acacctcccc
 1081 ctgaacctga aacataaaat gaatgcaatt gttgttgta
 acttgtttat tgcagcttat

1141 aatggttaca aataaagcaa tagcatcaca aatttcacaa
 ataaagcatt tttttcactg
 1201 cattctagtt gtggtttgtc caaactcatc aatgtatctt
 aacgcggatc tgggcgtggt
 1261 taaggggtggg aaagaatata taaggtgggg gtcttatgta
 gttttgtatc tgttttgca
 1321 cagccgcccgc cgccatgagc accaactcgt ttgatggaag
 cattgtgagc tcatatttga
 1381 caacgcgcat gccccatgg gccggggtgc gtcagaatgt
 gatgggctcc agcattgatg
 1441 gtgcgccggt cctgcccga aacttacta ccttgacct
 cgagaccgtg tctggaacgc
 1501 cgttggagac tgcagcctcc gcgcgcgctt cagcogctgc
 agccaccgcc cgcgggattg
 1561 tgactgactt tgctttcctg agcccgctt caagcagtgc
 agcttcccgt tcatccgcc
 1621 gcgatgacaa gttgacggct cttttggcac aattggattc
 tttgaccggg gaacttaatg
 1681 tegtctctca gcagctgttg gatctgcgcc agcaggttc
 tgcctgaag gcttctccc
 1741 ctcccaatgc ggtttaaacc ataaataaaa aaccagactc
 tgtttggatt tggatcaagc
 1801 aagtgtcttg ctgtctttat ttaggggtt tgccgcgcgcg
 gtaggcccgg gaccagcgg
 1861 ctcggtcgtt gagggtcctg tgtatttttt ccaggacgtg
 gtaaaggtga ctctggatgt
 1921 tcagatacat gggcataagc ccgtctctgg ggtggaggta
 gcaccactgc agagcttcat
 1981 gctgcgggggt ggtgtttag atgatccagt cgtagcagga
 gcgctgggcg tgggtgcctaa
 2041 aatgtcttt cagtagcaag ctgattgcca ggggcaggcc
 cttggtgtaa gtgtttacaa
 2101 agcggttaag ctgggatggg tgcatacgtg gggatatgag
 atgcatcttg gactgtattt
 2161 ttaggttggc tatgttccca gccatatccc tccggggatt
 catgttgtgc agaaccacca
 2221 gcacagtga tccgggtgcac ttgggaaatt tgtcatgtag
 cttagaagga aatgcgtgga
 2281 agaacttga gacgcccttg tgacctcaa gattttcoat
 gcattcgtcc ataatgatgg
 2341 caatgggccc acgggcggcg gcctgggcca agatatttct
 gggatcacta acgtcatagt
 2401 tgtgttccag gatgagatcg tcataggcca tttttacaaa
 gcgcgggccc agggtgccag
 2461 actgcggtat aatggttcca tccggcccag gggcgtagtt
 accctcacag attgcattt
 2521 cccacgcttt gagttcagat ggggggatca tgtctacctg
 cggggcgatg aagaaaacgg
 2581 tttccgggggt aggggagatc agctgggaag aaagcaggtt
 cctgagcagc tgcgacttac

2641 cgcagccggt gggcccgtaa atcacaccta ttaccgggtg
 caactggtag ttaagagagc
 2701 tgcagctgcc gtcacccctg agcagggggg ccaacttcgt
 aagcatgtcc ctgactcgca
 2761 tgttttccct gaccaaatcc gccagaaggc gctcgcgcc
 cagcgatagc agttcttgca
 2821 aggaagcaaa gtttttcaac ggtttgagac cgtccgcgct
 aggcattgctt ttgagcgctt
 2881 gaccaagcag ttccaggcgg tcccacagct cggtcacctg
 ctctacggca tctcgatcca
 2941 gcatatctcc tcgtttcgcg ggttggggcg gctttcgcgtg
 tacggcagta gtcggtgctc
 3001 gtccagacgg gccaggggtca tgtctttcca cgggagcagg
 gtccctcgtca gcgtagtctg
 3061 ggtcacgggtg aaggggtgcg ctccgggctg cgcgctggcc
 aggggtgctt tgaggctggt
 3121 cctgctgggtg ctgaagcgtt gccggctctt gccctgcgcg
 tcggccaggt agcatttgac
 3181 catgggtgtca tagtccagcc cctccgcggc gtggcccttg
 gcgagcagct tgcccttggc
 3241 ggaggcgccg cagcaggggc agtgcagact tttgagggcg
 tagagcttgg gcgagagaaa
 3301 taccgattcc ggggagtagg catccgcgcc gcaggccccg
 cagacggctt cgcattccac
 3361 gagccaggtg agctctggcc gttcgggggtc aaaaaccagg
 tttccccat gctttttgat
 3421 gcgtttctta cctctggttt ccatgagccg gtgtccacgc
 tcggtgacga aaaggctgtc
 3481 cgtgtccccg tatacagact tgagagggag tttaaacgaa
 ttcaatagct tgttgcatgg
 3541 gcggcgatat aaaatgcaag gtgctgctca aaaaatcagg
 caaagcctcg cgcaaaaaag
 3601 aaagcacatc gtagtcatgc tcatgcagat aaaggcaggt
 aagctccgga accaccacag
 3661 aaaaagacac catttttctc tcaaacatgt ctgcggggtt
 ctgcataaac acaaaataaa
 3721 ataacaaaaa aacatttaaa cattagaagc ctgtcttaca
 acaggaaaaa caacccttat
 3781 aagcataaga cggactacgg ccatgccggc gtgaccgtaa
 aaaaactggt caccgtgatt
 3841 aaaaagcacc accgacagct cctcgggtcat gtccggagtc
 ataattgtaag actcggtaaa
 3901 cacatcaggt tgattcatcg gtcagtgcta aaaagcgacc
 gaaatagccc gggggaatac
 3961 ataccgcag gcgtagagac aacattacag cccccatagg
 aggtataaca aattaatag
 4021 gagagaaaaa cacataaaca cctgaaaaac cctcctgcct
 aggcaaaata gcaccctccc
 4081 gctccagaac aacatacagc gcttcacagc ggcagcctaa
 cagtcagcct taccagtaaa

4141 aaagaaaacc tattaaaaa acaccactcg acacggcacc
 agctcaatca gtcacagtgt
 4201 aaaaaagggc caagtgcaga gcgagtatat ataggactaa
 aaaatgacgt aacggttaaa
 4261 gtccacaaaa aacaccacaga aaaccgcacg cgaacctacg
 cccagaaacg aaagccaaaa
 4321 aaccacaaac ttctcaaat cgtaacttcc gttttccac
 gttacgtaac ttcccatttt
 4381 aagaaaacta caattcccaa cacatacaag ttactccgcc
 ctaaaaccta cgtaaccgac
 4441 cccgttccca cgccccgacg cacgtcacia actccacccc
 ctcaattatca tattggcttc
 4501 aatccaaaat aaggtatatt attgatgatg ttaattaaca
 tgcattgatc catatgcggt
 4561 gtgaaatacc gcacagatgc gtaaggagaa aataccgcat
 caggcgctct tccgcttctt
 4621 cgctcaactga ctgctgacg tcggtcgttc ggctgcggcg
 agcggtatca gctcaactca
 4681 aggcggtaat acggttatcc acagaatcag gggataacgc
 aggaaagaac atgtgagcaa
 4741 aaggccagca aaaggccagc aaccgtaaaa aggccgcggt
 gctggcggtt ttccataggc
 4801 tccgcccccc tgacgagcat cacaaaaatc gacgctcaag
 tcagaggtgg cgaaaccgca
 4861 caggactata aagataccag gcgtttcccc ctggaagctc
 cctcgtgacg tctctgttc
 4921 cgaccctgcc gcttaccgga tacctgtccg cctttctccc
 ttccggaagc gtggcgcttt
 4981 ctcatagctc acgctgtagg tatctcagtt cgggtgtagg
 cgctcgtcc aagctgggct
 5041 gtgtgcacga accccccgct cagccccgacc gctgcgcctt
 atccggtaac tatcgtcttg
 5101 agtccaaccc ggtaagacac gacttatcgc cactggcagc
 agccactggt aacaggatta
 5161 gcagagcgag gtatgtaggc ggtgctacag agttcttgaa
 gtgggtggct aactacggct
 5221 aactagaag gacagtattt ggtatctgac ctctgctgaa
 gccagttacc ttcggaaaaa
 5281 gagttggtag ctcttgatcc ggcaaaaaa ccaccgctgg
 tagcgggtgt tttttgttt
 5341 gcaagcagca gattacgac agaaaaaaag gatctcaaga
 agatcctttg atcttttcta
 5401 cggggtctga cgctcagtgg aacgaaaact cacgttaagg
 gattttggct atgagattat
 5461 caaaaaggat cttcacctag atccttttaa attaaaaatg
 aagtttttaa tcaatctaaa
 5521 gtatatatga gtaaacttgg tctgacagtt accaatgctt
 aatcagtgag gcacctatct
 5581 cagcagctctg tctatttctg tcatccatag ttgcctgact
 ccccgtcgtg tagataacta

5641 cgatacggga gggcttacca tctggcccca gtgctgcaat
 gataccgcga gacccacgct
 5701 caccggctcc agatttatca gcaataaacc agccagccgg
 aagggccgag cgcagaagtg
 5761 gtcttgcaac tttatccgcc tccatccagt ctattaattg
 ttgccgggaa gctagagtaa
 5821 gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat
 tgctgcagcc atgagattat
 5881 caaaaaggat cttcacctag atccttttca cgtagaaagc
 cagtccgcag aaacggtgct
 5941 gaccccgat gaatgtcagc tactgggcta tctggacaag
 ggaaaacgca agcgcaaaga
 6001 gaaagcaggt agcttgcaat gggcttacat ggcgatagct
 agactgggcg gttttatgga
 6061 cagcaagcga accggaattg ccagctgggg cgccctctgg
 taaggttggg aagccctgca
 6121 aagtaaactg gatggcttcc tcgccgcaa ggatctgatg
 gcgcagggga tcaagctctg
 6181 atcaagagac aggatgagga tcgtttcgca tgattgaaca
 agatggattg cacgcaggtt
 6241 ctccggccgc ttgggtggag aggctattcg gctatgactg
 ggcacaacag acaatcggct
 6301 gctctgatgc cgccgtggtc cggctgtcag cgcaggggcg
 cccggttctt tttgtcaaga
 6361 ccgacctgtc cggtgccctg aatgaactgc aagacgaggc
 agcgcgggcta tcgtggctgg
 6421 ccacgacggg cgttccttgc gcagctgtgc tcgacgttgt
 cactgaagcg ggaaggact
 6481 ggctgctatt gggcgaagtg ccggggcagg atctcctgtc
 atctcacctt gctcctgcog
 6541 agaaagtatc catcatggct gatgcaatgc ggcggctgca
 tacgcttgat ccggctacct
 6601 gccattcga ccaccaagcg aaacatcgca tcgagcgagc
 acgtactcgg atggaagcog
 6661 gtcttgctga tcaggatgat ctggacgaag agcatcaggg
 gctcgcgcca gccgaactgt
 6721 tcgccaggct caaggcgagc atgcccagcg gcgaggatct
 cgctcgtgacc catggcgatg
 6781 cctgcttgcc gaatatcatg gtggaaaatg gccgcttttc
 tggattcatc gactgtggcc
 6841 ggctgggtgt ggcggaccgc tatcaggaca tagcgttggc
 taccctgat attgctgaag
 6901 agcttggcgg cgaatgggct gaccgcttcc tcgtgcttta
 cggtatcgcc gctcccatt
 6961 cgcagcgcac cgccttctat cgccttcttg acgagttctt
 ctgaattttg taaaatttt
 7021 tgttaaataca gctcattttt taaccaatag gccgaaatcg
 gcaacatccc ttataaatca
 7081 aaagaataga ccgcgatagg gttgagtgtt gttccagttt
 ggaacaagag tccactatta

7141 aagaacgtgg actccaacgt caaagggcga aaaaccgtct
 atcagggcga tggcccacta
 7201 cgtgaaccat cacccaaate aagttttttg cggtcgaggt
 gccgtaaagc tctaaatcgg
 7261 aaccctaaag ggagcccccg atttagagct tgacggggaa
 agccggcgaa cgtggcgaga
 7321 aaggaaggga agaaagcgaaggagcgggc gctagggcgc
 tggcaagtgt agcggtcacg
 7381 ctgcgcgtaa ccaccacacc cgcgcgctta atgocccgct
 acagggcgcg tccattcgc
 7441 attcaggatc gaattaattc ttaat

SEQ ID: 14 DS2beta-luc

taacatcatc aataatatac cttattttgg attgaagcca atatgataat
 gagggggtgg agtttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
 acgtagtagt gtggcggaag tgtgatggtg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgttt ttggtgtgcg ccggtgtaca
 caggaagtga caatttctgc gcggttttag gcggatggtg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtgaaa totgaataat tttgtgttac tcatagcgcg
 taatact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**
CAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
TAGGGCGAATTGGGTACTGGCCAcagagccttggcccattgcatacgttgtatccatata
atatgtacatttatattggctcatgtccaacattacgccatggtgacattgattattgact
agttattaatagtaatacaattacggggtcattagttcatagccatataatggagttccgcgt
tacataacttacggtaaatggcccgcctggctgaccgcccacgacccccgccattgacgt
caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
gagtatttacggtaaaactgcccacttggcagtacatcaagtgtatcata tgccaagtacgcc
ccattgacgtcaatgacggtaaatggcccgcctggcattatgccagtacatgacctat
gggactttcctacttggcagtacatcctacgtattagtcacgcctattaccatggtgatgcgg
ttttggcagtacatcaatgggctggatagcgggtttgactcaagggtttccaagtcca
ccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtc
gtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacgggtgggaggtctatata
agcagagctcgttttagtgaaccgtcagatcgcctggagacgccatccacgctgttttgacct
ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttggtggtgag
gccctgggcaggttggtatcaaggttacaagacaggtttaaggagaccaatagaaactgggc
atgtggagacagagaagactccttgggtttctgataggcaactgactctctctgcctattggtc
tattttcccacccttaggctgctggtctgagcctagGAGATCTCTCGAGGTCTGACGGTATCG
ATGGGTACCGCGCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGC
CTCGACTGTGCCTT
CTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
ACTCCCAC
TGTCCTTTCCTAATAAAATGAGGAAATGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
TGGGGGGT
GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
GGTGGGCT
CTATGGCTTCTGAGGCGGAAAGAACCTATGGCTTCTGAGGCGGAAAGAACCCA**ACCACCGCGG**
TGGCGGCCGCCACACAAAAACCAACACACAGATGTAATGAAAATAAAGATATTTTATTTCT
AGAGAAACGATATGGGCTGAATACGGATCCGTATTTCAGCCATATCGTTTCCTGCAGGAATT
CGCCCTTAGATATCATCGATGTCTCGGCGGTGGTGGCGCGTCGCGCCGCTGGGTTTTATAG

GGCGCCGCGCGGCCGCTCGAGCCATAAAAGGCAACTTTCGGAACGGCGCACGCTGATTGGC
 CCGCGCCGCTCACTCACCGGCTTCGCCGCACAGTGCAGCATTTTTTTACCCCTCTCCCCT
 CCTTTTGCAGAAAAAAGAGCGAGAGCGAGATTGAGGAAGAGGAGGAGGGAGAGTTTTG
 GCGTTGGCCGCTTGGGGTGCTGGGCGTCGACGATATCTAAGGGCGAATTCGATATCAagct
 agct**Tgtcgactcg** aagatctggg cgtgggtaag
 421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctggt ttgcagcagc
 481 cgccgcgcgc atgagcacca actcgtttga tgggaagcatt
 gtgagctcat atttgacaac
 541 gcgcatgccc ccatgggccc ggggtgcgtca gaatgtgatg
 ggetccagca ttgatggctg
 601 ccccgctctg cccgcaaac ctactacctt gacctacgag
 accgtgtctg gaacgcctt
 661 ggagactgca gcctccgccc cgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac
 721 tgactttgct ttctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga
 781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggaac ttaatgtcgt
 841 ttctcagcag ctggttgatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcgggt taaaacataa ataaaaaac agactctggt
 tggatttggg tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcggctctg
 1021 gtcgttgagg gtccctgtgta tttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccgt ctctgggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cgggggtggg ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggg gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttgact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tcactaacgt catagttgtg
 1561 ttccaggatg agatcgtcat aggccatttt taciaagcgc
 gggcggaggg tgccagactg
 1621 cggtataatg gttccatccg gccaggggc gtagttaccc
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct gggaagaaag caggttcctg
 agcagctgcg acttaccgca

1801 gccggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggccac ttcgttaagc
 atgtccctga ctgcatggt
 1921 tccctgacc aatccgcca gaaggcgctc gccgccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacgggt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcgggtccc acagctcggc cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttccgagggt ggggaggctt tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacgggcca gggtcatgtc tttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctcctg
 2281 ctgggtgctga agcgcctgcc gtcttcgccc tgcgcgtcgg
 ccaggtagca tttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcggtg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgccgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggccc gagaaatacc
 2461 gattccgggg agtaggcctc cgcgccgcag gccccgcaga
 cggctctcga ttccacgagc
 2521 caggtgagct ctggccggtc ggggtcaaaa accaggtttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagtta aacgaattca
 atagcttggt gcatgggccc
 2701 cgatataaaa tgcaagggtc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaaccatt agaaagcctgt cttacaacag
 gaaaaacaac cttataagc
 2941 ataagacgga ctacggccat gccggcggtga ccgtaaaaaa
 actggtcacc gtgattaata
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcgggtc gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgtc
 3241 cagaacaaca tacagcgtt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag

3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca cccagaaaac cgcacgcgaa cctacgceca
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaatcgtc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tcccaacaca tacaagttac tccgccctaa
 aacctagctc acccgccccg
 3601 ttcccacgcc cgcgcgccag tcacaaactc cccccctca
 ttatcatatt ggcttcaatc
 3661 caaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg cttcctcgct
 3781 cactgactcg ctgcgctcgg tcggtcggct gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcggtgctg
 gcgtttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcggg gtaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtgggtttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaacg aaaactcacg ttaagggatt
 ttggtcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcggtgtaga taactacgat

4801 acgggagggc ttaccatctg gcccagtgcc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgcga gaagtggctc
 4921 tgcaacttta tccgcctcca tccagtctat taattgttgc
 cgggaagcta gagtaagtag
 4981 ttgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccgatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt ttccatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgcc gtgttccggc tgtcagcgca ggggcgcccg
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 ccgctatcgt ggctggccac
 5581 gacgggcggt ccttgccgag ctgtgctcga cgttgctact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgccgagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatacga gcgagcacgt
 actcggatgg aagccggtct
 5821 tgtcgcacag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gcgatgcctg
 5941 cttgccgaat atcatggtgg aaaatggccg cttttctgga
 ttcacgcact gtggccggct
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttctcgt gctttacggt
 atcgcgcctc ccgattcgca
 6121 gcgcatacgc ttctatcgc ttcttgacga gttcttctga
 attttgttaa aatttttgtt
 6181 aatcagctc attttttaac caataggccg aatcggcaa
 catcccttat aatcaaaaag
 6241 aatagaccgc gatagggttg agtggtgttc cagtttgaa
 caagagtcca ctattaaaga

6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccaactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgaggtgccg
 taaagctota aatcggaacc
 6421 ctaaaggag cccccgattt agagcttgac ggggaaagcc
 ggcgaacgtg gcgagaaagg
 6481 aagggaaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgcgtcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID: 15 taacatcatic aataatatac cttattttgg attgaagcca
 atatgataat gagggggtg agtttgtgac gtggcgcggg gcgtgggaac
 gggcggggtg acgtagtagt gtggcggaag tgtgatggtg caagtgtggc
 ggaacacatg taagcgacgg atgtggcaaa agtgacgttt ttggtgtgcy
 ccggtgtaca caggaagtga caattttcgc gcggttttag gcggatggtg
 tagtaaattt gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa
 ctgaataaga ggaagtgaat tctgaataat tttgtgttac tcatagcgcg
 taatact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**
CAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
TAGGGCGAATTGGGTACTGGCCAacagagcttggccccattgcatacgttgtatccatatacata
 atatgtacatttatattggctcatgtccaacattaccgccatggtgacattgattattgact
 agttattaatagtaatcaattacggggctcattagttcatagcccatatattggagttccgcgt
 tacataacttacggtaaatggccccctggctgaccgccaacgacccccgccattgacgt
 caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtatttacggtaaactgccacttggcagtaacatcaagtgtatcatatgccaaagtacgcc
 cctattgacgtcaatgacggtaaatggccccctggcattatgccagtaacatgacctat
 gggactttcctacttggcagtaacatca**cgtaattagtcacgctatttaoccat**ggtgatgcgg
 ttttggcagtaacatcaatgggcggtgtagcgggttgactcacggggatttccaagtctcca
 cccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtc
 gtaacaaactccgccccattgacgcaaatgggcggttaggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgctcagatcgctggagacgccatccacgctgttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttggtggtgag
 gccctgggcaggttggatcaaggttacaagacaggtttaaggagaçcaatagaaact**gggc**
atgtggagacagagaagactcttgggtttctgataggcactgactctctctgcctattggtc
 tattttccacccttaggctgctggtctgag**cctagGAGATCTCTCGAGGTCGACGGTATCG**
ATGGGTACCGCGGCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGC
CTCGACTGTGCCTT
 CTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCAC
 TGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
 TGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGCT
 CTATGGCTTCTGAGGCGGAAAGAACCA

AAGCTTAGGCTCGAGCGGCC
 GCCACACAAAAACCAACACACAGATGTAATGAAAATAAAGATATTTTATTTCTAGAGAAAC
 GATATGGGCTGAATACGGATCCGTATTTCAGCCATATCGTTTCCAGATCTCTAGCGGATCT
 GACGGTCACTAAACCAGCTCTGCT

TATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTT
 ACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCATTGACGTCAAT
 GGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTTGATGTACTG
 CAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGG
 AAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGAC
 GTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTA
 CCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAAC
 ATACGTCAATTATTGACGTCAATGGGCGGGGGTCTGTTGGGCGGTGAGCCAGGCGGGCCATT
 TACCGTAAGTTATGTAACGCGGAACTCCATATATGGGCTATGAACTAATGACCCCGTAAT
 TGATTACTATTACAGTATTACGCGCTATGAGTAACACAAAATTATTTCAGATTTCACTTCC
 TCTTATTTCAGTTTTTCCCGCGAAAATGGCCAAATCTTACTCGGTTACGCCCAAATTTACTA
 CAACATCCGCCTAAAACCGCGCGAAAATTGTCACTTCCCTGTGTACACCGGCGCACACCAA
 AAACGTCACTTTTGGCCACATCCGTGCTTACATGTGTTCCGCCACACTTGCAACATCACA
 CTTCCGCCACACTACTACGTCAACCGCCCCGTTCCACGCCCCGCGCCACGTCAAAAAT
 CCACCCCTCATTATCATATTGGCTTCAATCCAAAATAAGGTATATTATTGATGATGTTAAG
 CTT

gtcgactcg aagatctggg cgtggttaag
 421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctggt ttgcagcagc
 481 cgccgcgccc atgagcacca actcgtttga tgggaagcatt
 gtgagctcat atttgacaac
 541 gcgcatgccc ccattgggccc ggggtgcgtca gaatgtgatg
 ggctccagca ttgatggctg
 601 ccccgctctg cccgcaaact ctactacctt gacctacgag
 accgtgtctg gaacgcccgtt
 661 ggagactgca gcctccgccc ccgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac
 721 tgactttgct ttccctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga
 781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccgggaac ttaatgtcgt
 841 ttctcagcag ctgtttggatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcgggt taaaacataa ataaaaaacc agactctggt
 tggatttggga tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcgggtctg
 1021 gtcgttgagg gtccctgtgta ttttttcag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccgt ctctggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cgggggtggtg ttgtagatga tccagtcgta gcaggagcgc
 tgggctggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttgact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa

1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tcactaacgt catagtgtg
 1561 ttccaggatg agatcgatc aggccatttt tacaagcgc
 gggcggaggg tgccagactg
 1621 cggataaatg gttccatccg gcccaggggc gtagttacc
 tcacagattt gcatttcca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggttc
 1741 cggggtaggg gagatcagct gggaagaaag caggttcctg
 agcagctgcg acttaccgca
 1801 gccgggtggc ccgtaaataca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aaatccgcca gaaggcgtc gccgccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacggtt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcgggtccc acagctcggg cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttcgcgggtt ggggcggctt tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacgggcca gggatcatgtc tttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctcctg
 2281 ctggtgctga agcgtgcgcg gtcttcgccc tgcgctcgg
 ccaggtagca tttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgccgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggccc gagaaatacc
 2461 gattccgggg agtaggcata cgcgccgcag gccccgcaga
 cggctctgca ttccacgagc
 2521 caggtgagct ctggccgttc ggggtcaaaa accaggtttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagttaa aacgaattca
 atagcttggt gcatgggcgg
 2701 cgatataaaa tgcaaggcgc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac ccttataagc

2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaana
 3001 agcaccaccg acagctcctc ggatcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcgggtca gtgctaanaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgacggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgtc
 3241 cagaacaaca tacagcgctt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtanaa
 3361 aaggccaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 aaaaaaaca cccagaaaac cgcacgcgaa cctacgcca
 gaaacgaaag ccaaaaaacc
 3481 cacaacttc tcaaatcgtc acttccgttt tcccacgta
 cgtaacttc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgcctaa
 aacctacgtc acccgccccg
 3601 tccccacgc ccgcccacg tcacaaactc caccacctca
 ttatcatatt ggcttcaatc
 3661 caaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcgggtgga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttcg ctctctcgt
 3781 cactgactcg ctgcgctcgg tcgctcggct gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt tccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc ttaggtatc tcagttcggg ttaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgcttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggtat gtaggcgggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt

4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtgggtttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 ctttgatct tttctacggg
 4561 gtctgacgct cagtggaacg aaaactcacg ttaagggatt
 ttggatcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttggctctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta ttcggttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gccccagtgc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgcga gaagtggctc
 4921 tgcaacttta tccgcctcca tccagtctat taattgttgc
 cgggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaaac ggtgctgacc
 5101 ccgatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt ttcgcatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgcc gtgttccggc tgtcagcgcga ggggccccg
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggcgtt ccttgcgag ctgtgctega cgttgtcact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgccgagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcacga gcgagcacgt
 actcggatgg aagccggtct
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacctatg gcgatgcctg

5941 cttgccgaat atcatgggtgg aaaatggccg cttttctgga
 ttcategact gtggccggct
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttcctcgt gctttacggt
 atcgccgctc ccgattcgca
 6121 ggcategcc ttctatgcc ttcttgacga gttcttctga
 attttgtaa aattttgtt
 6181 aatcagctc atttttaac caataggccg aaatcggcaa
 catcccttat aaatcaaaag
 6241 aatagaccgc gatagggttg agtggtgttc cagtttgaa
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc cactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgagggtgccg
 taaagctcta aatcggaacc
 6421 ctaaaggag cccccgattt agagcttgac ggggaaagcc
 ggcgaacgtg gcgagaaagg
 6481 aaggaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgctcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID: 16 vector lanzadera HA, CMV-luc

taacatcacc aataatatac cttattttgg attgaagcca atatgataat
 gaggggggtgg agtttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
 acgtagtagt gtggcggaag tgtgatgttg caagtgtggc ggaacacatg
 taagcagcgg atgtggcaaa agtgacgttt ttgggtgtgcg ccgggtaca
 caggaagtga caatttcgc gcggtttttag gcggatgttg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtgaaa totgaataat tttgtgttac tcatagcgcg
 taataact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**
CAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
TAGGGCGAATTGGGTACTGGCCAacagagcttggcccattgcatacgttgtatccatatcata
 atatgtacatttatattggctcatgtccaacattaccgccatgttgacattgattatt**gact**
agttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttccgcgt
 tacataacttacggtaaatggcccgcctggctgaccgccaacgacccccgccattgacgt
 caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtatttacggtaaacgcccacttggcagtacatcaagtgtatcatatgccaaagtacgcc
 ccctattgacgtcaatgacggtaaatggcccgcctggcattatgccagtacatgacctat
 gggactttccctacttggcagtacatct**agttattagtcacgcctattaccat**gggtgatgagg
 ttttggcagtacatcaatgggcgtggatagcgggttgactcacggggatttccaagtctcca
 ccccatgacgtcaatgggagtttgttttggcaccâaaatcaacgggactttccaaaatgtc
 gtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacgggtgggaggtctatata
 agcagagctcgtttagtgaaaccgtcagatcgccctggagacgccatccacgctgttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttgggtggtgag
 gccctgggcaggttggatcaaggttacaagacaggtttaaaggagaccaatagaaact**gggc**
atgtggagacagagaagactcttgggtttctgataggcaactgactctctctgcctattggctc

tat tttccacccttaggctgctggtctgagcctag**GAGATCTCTCGAGGTCGACGGTAT**cg
atgccca ccATGGAGAA AATCGTCCTG TTGCTCGCTA TTGTGTCTCT

51 AGTGAAGAGC GATCAAATTT GTATCGGCTA CCATGCCAAT AACTCAACAG
101 AGCAGGTCTGA TACTATCATG GAGAAAAACG TAACAGTTAC TCATGCCCAA
151 GACATCTTGG AAAAGACCCA CAACGGCAAA CTTTGGCGACC TGGATGGAGT
201 GAAGCCCCTG ATCCTCCGGG ACTGTTTCAGT CGCTGGTTGG CTGCTCGGGA
251 ACCCTATGTG TGATGAGTTT ATCAACGTGC CTGAATGGTC TTACATTGTG
301 GAGAAGGCTA ACCCTACCAA TGACCTCTGC TATCCTGGGT CATTTAACGA
351 TTACGAGGAA CTGAAACACC TGTTGTCTAG AATTAACCAC TTTGAAAAGA
401 TACAGATTAT ACCCAAGTCT AGTTGGAGTG ATCACGAAGC CTCCTCAGGC
451 GTTAGCTCAG CGTGTCCCTA TCTGGGCTCT CCATCCTTCT TTAGAAATGT
501 GGTCTGGTTA ATCAAAAAGA ACAGTACCTA CCAACCATC AAAAAGTCTT
551 ATAACAATAC CAATCAGGAG GACCTGCTCG TGTTGTGGGG TATCCATCAC
601 CCGAACGACG CCGCTGAACA GACTAGCTG TATCAGAACC CCAATACATA
651 CATCAGTATT GGCACGAGTA CTCTGAACCA GCGATTAGTG CCAAAGATTG
701 CAACACGGAG CAAAGTAAAT GGGCAATCTG GCAGGATGGA GTTTTTCTGG
751 ACAATCTTAA AACCCAACGA TGCATAAAT TTCGAGTCCA ATGGCAATTT
801 CATCGCCCCT GAATACGCCT ATAAGATCGT GAAAAAGGGG GACTCTGCAA
851 TTATGAAGTC CGAATTAGAG TATGGCAATT GCAACACGAA GTGCCAGACA
901 CCAATGGGAG CCATTAATAG CTCAATGCC TTCCATAATA TTCATCCATT
951 GACCATTGGG GAGTGCCCAA AGTACGTGAA GTCCAACCGC CTGGTCCCTCG
1001 CAACCGGTCT AAGAAATAGC CCGCAGAGAG AATCGCGGAG GAAGAAACGT
1051 GGCCTGTTTG GCGCGATTGC CGGATTCATC GAGGGAGGCT GGCAGGGTAT
1101 GGTCGATGGT TGGTACGGAT ACCACCATAG CAACGAACAG GGGTCCGGCT
1151 ATGCAGCAGA TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAAC
1201 AAGGTTAATA GTATTATAGA TAAATGAAC ACGCAATTCG AGGCCGTGG
1251 GAGGGAGTTT AACAATCTGG AACGCCGGAT CGAAAATCTG AATAAGAAAA
1301 TGGAAGACGG CTTCCTTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA
1351 CTCATGGAGA ACGAGAGGAC CCTGGATTC CACGATAGCA ACGTGAAGAA
1401 CCTTTACGAC AAGGTGAGAC TTCAGCTCCG AGACAACGCC AAGGAGCTGG
1451 GGAATGGATG CTTGAGTTT TACCACAAAT GTGACAATGA GTGCATGGAA
1501 AGTATACGCA ACGGGACCTA CAATTACCCT CAGTATAGCG AAGAGGCTCG
1551 GCTCAAACGC GAAGAGATAA GCGGGGTGAA ATTGGAATCA ATCGGAACAT
1601 ATCAAATCCT GTCCATCTAT TCCACCGTCG CCTCTTCGCT GGCCCTCGCT
1651 ATCATGATGG CTGGTCTGTC CCTATGGATG TGTTCCAATG GAAGCCTTCA
1701 GTGCCGTATT TGTATATGAg c

GGCCGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCC****
TT

CTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
ACTCCCAC

TGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
TGGGGGGT

GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
GGTGGGCT

CTATGGCTTCTGAGGCGGAAAGA**CCAAAGCT**taacatcatc aataatatac
cttattttgg attgaagcca atatgataat gaggggggtgg

61 agtttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
acgtagtagt gtggcggaag

121 tgtgatgttg caagtgtggc ggaacacatg taagcgacgg
atgtggcaaa agtgacgttt

181 ttgggtgtgcg cgggtgtaca caggaagtga caattttcgc
 gcggtttttag gcggatggtg
 241 tagtaaattt gggcgtaacc gagtaagatt tggccatttt
 cgcgggaaaa ctgaataaga
 301 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg
 taatactgta atagtaatca
 361 attacggggt cattagttca tagcccatat atggagttcc
 gcgttacata acttacggta
 421 aatggcccgc ctggctgacc gcccaacgac ccccgcccat
 tgacgtcaat aatgacgtat
 481 gttcccatag taacgccaat agggactttc cattgacgtc
 aatgggtgga gtatttacgg
 541 taaactgccc acttggcagt acatcaagtg tatcatatgc
 caagtacgcc ccctattgac
 601 gtcaatgacg gtaaattggcc cgcctggcat tatgcccagt
 acatgacctt atgggacttt
 661 cctacttggc agtacatcta cgtattagtc atcgctatta
 ccatggtgat gcggtttttg
 721 cagtacatca atgggcgtgg atagcggttt gactcacggg
 gatttccaag tctccacccc
 781 attgacgtca atgggagttt gttttggcac caaaatcaac
 gggactttcc aaaatgtcgt
 841 aacaactccg cccattgac gcaaatgggc ggtaggcgtg
 tacggtgggg ggtctatata
 901 agcagagctg gtttagtgaa cgcgcagatc
 cgctag**AGATCTGG**gaaacgatatgggctgaatac**ggatccg**tattcagcccatatcgtttc
T CT AGA AAT AAA ATA TCT TTA TTT TCA TTA CAT CTG TGT GTT GGT
 TTT TTG TGT GGC
GG CCGCctagacc **AAGCTT**ctag ataagatatc cgatccaccg gatctagata
 actgatcata
 1021 atcagccata ccacatttgt agaggtttta cttgctttaa
 aaaacctccc acacctcccc
 1081 ctgaacctga aacataaaat gaatgcaatt gttgttgta
 acttgtttat tgcagcttat
 1141 aatggttaca aataaagcaa tagcatcaca aatttcacaa
 ataaagcatt ttttcaactg
 1201 cattetagt gtggtttgtc caaactcacc aatgtatctt
 aacgcggatc tgggcgtggt
 1261 taaggtggg aaagaatata taaggtggg gtcttatgta
 gttttgtatc **tgttttgcag**
 1321 **cagccgccc** cgccatgagc accaactcgt ttgatggaag cattgtg
AGCTTgtcgactcg aagatctggg cgtggttaag
 421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctggt ttgcagcagc
 481 cgccgcccgc atgagcacca actcgtttga tggaagcatt
 gtgagotcat atttgacaac
 541 gcgcatgccc ccatgggccc ggggtcgtca gaatgtgatg
 ggctcoagca ttgatggtcg
 601 cccgctcctg cccgcaaact ctactacctt gacctacgag
 accgtgtctg gaacgccgtt

661 ggagactgca gcctccgccg cgccttcagc cgctgcagcc
 accgcccgcg ggattgtgac
 721 tgactttgct ttcctgagcc cgccttgcaag cagtgcaagct
 tcccgttcat ccgcccgcga
 781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggaac ttaatgtcgt
 841 ttctcagcag ctgttgatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcggtt taaaacataa ataaaaaacc agactctggt
 tggatttggga tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcggctcgc
 1021 gtcgttgagg gtccctgtgta tttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccgt ctctggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cggggtggtg ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttggact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tcactaacgt catagtgtg
 1561 ttccaggatg agatcgtcat aggccatttt taciaagcgc
 gggcggaggg tgccagactg
 1621 cgggtataatg gttccatccg gccaggggc gtagttacce
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gogatgaaga aaacggttc
 1741 cggggtaggg gagatcagct gggagaagaa caggttctctg
 agcagctgcg acttaccgca
 1801 gccggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aatccgcca gaaggcgtc gccgccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacggtt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcggtccc acagctcggg cacctgctct
 acggcatctc gatccagcat
 2101 atctctcgt ttcgcgggtt gggcggcgtt tcgctgtacg
 gcagtagtgc gtgctcgtcc

2161 agacgggcca gggtcatgtc tttccacggg cgcagggctc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctctg
 2281 ctggtgctga agcgcctgcc gtcttcgccc tgcgcgtcgg
 ccaggtagca tttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgccgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggcgc gagaaatacc
 2461 gattccgggg agtaggcata cgcgccgcag gccccgcaga
 cggctctcga ttccacgagc
 2521 caggtgagct ctggccgttc ggggtcaaaa accaggtttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagtta aacgaattca
 atagcttggt gcatgggcgg
 2701 cgatataaaa tgcaagggtc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac cttataagc
 2941 ataagacgga ctacggccat gcggcgtga ccgtaaaaaa
 actggtcacc gtgattaata
 3001 agcaccaccg acagctctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcgggtc gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgtc
 3241 cagaacaaca tacagcgtt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca cccagaaaac cgcacgcgaa cctacgcccc
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaatcgtc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgccctaa
 aacctagtc acccgccccg
 3601 ttcccacgcc ccgcgccacg tcacaaactc caccacctca
 ttatcatatt ggcttcaatc

3661 caaaataagg tatattattg atgatgtaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg ctctctcgct
 3781 cactgactcg ctgcgctcgg tcgttcggct gcggcgagecg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcggtgctg
 gcgtttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgcccgtt accggatacc tgtccgcctt tctcccttcc
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcggg gtaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcc gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtgggtttt ttgtttcaa
 4501 gcagcagatt acgcgagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaacg aaaactcacg ttaagggatt
 ttggatcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttggctcg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gccccagtgc tgcaatgata
 ccgagagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtggctc
 4921 tgcaacttta tccgctcca tccagtctat taattggtgc
 cgggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccggatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa

5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt .ttcgcgatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgcc gtgttccggc tgtcagcgcga ggggcgcccg
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggcggt ccttgccgag ctgtgctcga cgttgctact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgccgagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatacga gcgagcacgt
 actcggatgg aagccggtct
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gcgatgcctg
 5941 cttgccgaat atcatgggtg aaaatggccg cttttctgga
 ttcacgcact gtggccggct
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttctcgt gctttacggg
 atcgcgcctc ccgattcgca
 6121 gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga
 attttgttaa aaTTTTTgTT
 6181 aaatcagctc attttttaac caataggccg aaatcggcaa
 catcccttat aaatcaaaag
 6241 aatagaccgc gatagggttg agtgttgttc cagtttgaa
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc cactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgagggtgccg
 taaagctcta aatcggaacc
 6421 ctaaaggag cccccgattt agagcttgac ggggaaagcc
 ggcaacgtg gcgagaaagg
 6481 aagggaagaa agcgaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacacccgcg cgcttaatgc gccgctacag
 ggcgctcca ttcgccatc
 6601 aggatcgaat taattcttaa t

1 taaggatccn nncctgtect cgaccgatgc ccttgagagc cttcaacca
 gtcagctcct
 61 tccggtgggc gcggggcatg actatcgtcg ccgcacttat
 gactgtcttc tttatcatgc
 121 aactcgtagg acaggtgccg gcagcgcctct gggtcatttt
 cggcgaggac cgctttcgct
 181 ggagcgcgac gatgatcggc ctgtcgccttg cggatttcgg
 aatcttgcac gccctcgcctc
 241 aagccttcgt cactggtccc gccaccaaac gtttcggcga
 gaagcaggcc attatcgccg
 301 gcatggcggc cgacgcgctg ggctacgtct tgctggcggt
 cgcgacgcga ggctggatgg
 361 ccttccccat tatgattctt ctcgcttccg gcggcatcgg
 gatgcccgcg ttgcaggcca
 421 tgctgtccag gcaggtagat gacgaccatc agggacagct
 tcaaggatcg ctgcggctc
 481 ttaccagcct aacttcgatc actggaccgc tgatcgtcac
 ggcgatttat gccgcctcgg
 541 cgagcacatg gaacggggtg gcatggattg taggcgccgc
 cctatacctt gtctgcctcc
 601 ccgcgttgcg tcgcggtgca tggagccggg ccacctcgac
 ctgaatggaa gccggcggca
 661 cctcgcctaac ggattcacca ctccaagaat tggagccaat
 caattcttgc ggagaactgt
 721 gaatgcgcaa accaaccctt ggcagaacat atccatcgcg
 tccgccatct ccagcagccg
 781 cacgcggcgc atctcgggca gcgttgggtc ctggccacgg
 gtgcgcatga tcgtgctcct
 841 gtcgttgagg acccggctag gctggcgggg ttgccttact
 ggttagcaga atgaatcacc
 901 gatacgcgag cgaacgtgaa gcgactgctg ctgcaaaacg
 tctgcgacct gagcaacaac
 961 atgaatggtc ttcggtttcc gtgtttcgta aagtctggaa
 acgcggaagt cagcgcctg
 1021 caccattatg ttccggatct gcatcgcagg atgctgctgg
 ctaccctgtg gaacacctac
 1081 atctgtatta acgaagcgtt ggcattgacc ctgagtgatt
 tttctctggg cccgccgcat
 1141 ccataccgcc agttgtttac cctcacaacg ttccagtaac
 cgggcatggt catcatcagt
 1201 aaccggtatc gtgagcatcc tctctcgttt catcggtatc
 attaccccca tgaacagaaa
 1261 ttccccctta cacggaggca tcaagtgacc aaacaggaaa
 aaaccgccct taacatggcc
 1321 cgctttatec gaagccagac attaacgctt ctggagaaac
 tcaacgagct ggacgcggat
 1381 gaacaggcag acatctgtga atcgcttcac gaccacgctg
 atgagcttta ccgcagctgc
 1441 ctgcgcggtt tcggatgatga cggtgaaaac ctctgacaca
 tgcagctccc ggagacggtc

ES 2 573 456 T3

1501 acagcttgtc tgtaagcgga tgccgggagc agacaagccc
 gtcagggcgc gtcagcgggt
 1561 gttggcgggt gtcggggcgc agccatgacc cagtcacgta
 gcgatagcgg agtgataact
 1621 ggcttaacta tgcggcatca gagcagattg tactgagagt
 gcaccatattg cgggtgtgaaa
 1681 taccgcacag atgcgtaagg agaaaatacc gcatcaggcg
 ctcttccgct tctctgctca
 1741 ctgactcgtc gcgctcggtc gttcggctgc ggcgagcggg
 atcagctcac tcaaaggcgg
 1801 taatacgggt atccacagaa tcaggggata acgcaggaaa
 gaacatgtga gcaaaaaggcc
 1861 agcaaaaaggc caggaaccgt aaaaaggccg cgttgctggc
 gtttttccat aggtcctcgc
 1921 cccctgacga gcatcacaaa aatcgacgct caagtcagag
 gtggcgaaaac ccgacaggac
 1981 tataaagata ccaggcgttt cccctggaa gctccctcgt
 gcgctctcct gttccgaccc
 2041 tgccgcttac cggatacctg tccgcctttc tcccttcggg
 aagcgtggcg ctttctcaat
 2101 gctcacgctg taggtatctc agttcgggtg aggtcgttcg
 ctccaagctg ggctgtgtgc
 2161 acgaaccccc cgttcagccc gaccgctgcg ccttatccgg
 taactatcgt cttgagtcca
 2221 acccggtaag acacgactta tcgccactgg cagcagccac
 tggtaacagg attagcagag
 2281 cgaggatgt aggcgggtgct acagagttct tgaagtggg
 gcctaactac ggctacacta
 2341 gaaggacagt atttggtatc tgcgctctgc tgaagccagt
 taccttcgga aaaagagttg
 2401 gtagctcttg atccggcaaa caaaccaccg ctggtagcgg
 tggttttttt gtttgcaagc
 2461 agcagattac gcgcagaaaa aaaggatctc aagaagatcc
 tttgatcttt tctacggggg
 2521 ctgacgctca gtggaacgaa aactcacggt aagggatttt
 ggtcatgaga ttatcaaaaa
 2581 ggatcttcac ctagatcctt ttaaattaa aatgaagttt
 taaatcaatc taaagtatat
 2641 atgagtaaac ttggtctgac agttaccaat gcttaatcag
 tgaggcacct atctcagcga
 2701 tctgtctatt tcgctcatcc atagttgctt gactccccgt
 cgtgtagata actacgatac
 2761 gggaggcgtt accatctggc ccagtgctg caatgatacc
 gcgagacca cgctcaccgg
 2821 ctccagattt atcagcaata aaccagccag ccggaagggc
 cgagcgcaga agtggctcctg
 2881 caactttatc cgcctccatc cagtctatta attggtgccc
 ggaagctaga gtaagtagtt
 2941 cgccagttaa tagtttgccg aacgttggtg ccattgctgc agccatga
 gattatcaaa

aaggatcttc acctagatcc ttttcacgta gaaagccagt cgcagaaac
ggtgctgacc
cggatgaat gtcagctact gggctatctg gacaagggaa aacgcaagcg
caaagagaaa
gcaggtagct tgcagtgggc ttacatggcg atagctagac tgggcggttt
tatggacagc
aagcgaaccg gaattgccag ctggggcgcc ctctggtaag gttgggaagc
cctgcaaagt
aaactggatg gctttctcgc cgccaaggat ctgatggcgc aggggatcaa
gctctgatca
agagacagga tgaggatcgt ttcgcatgat tgaacaagat ggattgcacg
caggttctcc
ggcgccttgg gtggagagggc tattcggcta tgactgggca caacagacaa
tcggctgctc
tgatgccgcc gtgttccggc tgtcagcgca ggggcgccc gttctttttg
tcaagaccga
cctgtccggt gccctgaatg aactgcaaga cgaggcagcg cggctatcgt
ggctggccac
gacgggcggt ccttgccgag ctgtgctcga cgttgtcact gaagcgggaa
gggactggct
gctattgggc gaagtgcggg ggcaggatct cctgtcatct caccttgctc
ctgccgagaa
agtatccatc atggctgatg caatgcggcg gctgcatacg cttgatccgg
ctacctgccc
attcgaccac caagcgaaac atcgcatcga gcgagcacgt actcggatgg
aagccggtct
tgtcgatcag gatgatctgg acgaagagca tcaggggctc gcgccagccg
aactgttgc
caggtcaag gcgagcatgc ccgacggcga ggatctcgtc gtgacctatg
gcgatgcctg
cttgccgaat atcatggtgg aaaatggccg cttttctgga ttcatcgact
gtggccggct
gggtgtggcg gacogctatc aggacatagc gttggctacc cgtgatattg
ctgaagagct
tggcggcgaa tgggctgacc gcttctcgt gctttacggg atcgccgctc
ccgattcgca
gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga attttgtaa
aatttttgtt
aatcagctc attttttaac caataggccg aatcggcaa catcccttat
aatcaaaaag
aatagaccgc gatagggttg agtgttgctc cagtttgaa caagagtcca
ctattaaaga
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc
ccactacgtg
aaccatcacc caaatcaagt tttttgcggt cgaggtgccg taaagctcta
aatcggaaacc
ctaaagggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg
gcgagaaaag
aagggaaagaa agcgaagga gcgggcgcta gggcgctggc aagtgtagcg
gtcacgctgc

gcgtaaccac cacacccgcg cgcttaatgc gccgctacag ggcgcgtcca
ttcggccattc

aggatcgaat taattcttaa ttaacatcatc aataatatac cttatcttgg
attgaagcca atatgataat gaggggggtg agtttgtgac gtggcgcggg
gcgtgggaac gggcggggtg acgtagtagt gtggcggaag tgtgatgtg
caagtgtggc ggaacacatg taagcgacgg atgtggcaaa agtgacgtt
ttgggtgtgcg ccggtgtaca caggaagtga caattttcgc gcggttttag
gcggatgttg tagtaaattt gggcgtaacc gagtaagatt tggccattt
cgcgggaaaa ctgaataaga ggaagtgaaa tctgaataat tttgtgttac
tcatagcgcg

taatact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAAACGC**
CAGGGTTTTCCCAGTCACGACGTTGTA~~AAA~~ACGACGGCCAGTGAATTGTAATACGACTCACTA
TAGGGCGAATTGGGTACTGGCCAcagagcttggcccattgcatacgttgtatccatata
atatgtacatttatattggctcatgtccaacattaccgccatgttgacattgattattgact
agttattaatagtaatcaattacgggggtcattagttcatagcccatatatggagttccgcgt
tacataacttacggtaaatggcccgcctggctgaccgccaacgacccccgccattgacgt
caataatgacgtatgttcccatagtaacgccaatagggacttccattgacgtcaatgggtg
gagtatttacggtaaactgccacttggcagtacatcaagtgtatcatatgccaaagtacgcc
ccctattgacgtcaatgacggtaaatggcccgcctggcattatgccagtagacatgacctat
gggactttcctacttggcagtacatcta**cg**tattag**tc**at**cg**ctatt**acc**atgggtgatgcgg
ttttggcagtacatcaatgggcgtggatagcggttgactcacggggatttccaagtctcca
ccccattgacgtcaatgggagtttggtttggcaccaaaatcaacgggacttccaaaatgct
gtaacaactccgccccattgacgcaaatgggcggttaggcgtgtacgggtgggaggtctatata
agcagagctcgttttagtgaaccgtcagatcgctggagacgccatccacgctgttttgacct
ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttgggtggtag
gccctgggcaggttgggtatcaaggttacaagacaggtttaaaggagaccaatagaaact**gggc**
atgtggagacagagaagactcttgggttctgataggeactgactctctctgcctattggtc
tattttcccacccttaggctgctggtctgag**cctagGAGATCTCTCGAGGTCGACGGTAT**cg
atgcca ccATGGAGAA AATCGTCCTG TTGCTCGCTA TTGTGTCTCTAGTGAAGAGC
GATCAAATTT GTATCGGCTA CCATGCCAAT AACTCAACAGAGCAGGTCGA
TACTATCATG GAGAAAAACG TAACAGTTAC TCATGCCCAA

GACATCTTGG	AAAAGACCCA	CAACGGCAAA	CTTTGCGACC	TGGATGGAGT
GAAGCCCCTG	ATCCTCCGGG	ACTGTTCACT	CGCTGGTTGG	CTGCTCGGGA
ACCCTATGTG	TGATGAGTTT	ATCAACGTGC	CTGAATGGTC	TTACATTGTG
GAGAAGGCTA	ACCCTACCAA	TGACCTCTGC	TATCCTGGGT	CATTTAACGA
TTACGAGGAA	CTGAAACACC	TGTTGTCTAG	AATTAACCAC	TTTGAAAAGA
TACAGATTAT	ACCCAAGTCT	AGTTGGAGTG	ATCACGAAGC	CTCCTCAGGC
GTTAGCTCAG	CGTGTCCCTA	TCTGGGCTCT	CCATCCTTCT	TTAGAAATGT
GGTCTGGTTA	ATCAAAAAGA	ACAGTACCTA	CCCAACCATC	AAAAAGTCTT
ATAACAATAC	CAATCAGGAG	GACCTGCTCG	TGTTGTGGGG	TATCCATCAC
CCGAACGACG	CCGCTGAACA	GACTAGGCTG	TATCAGAACC	CCACTACATA
CATCAGTATT	GGCACGAGTA	CTCTGAACCA	GCGATTAGTG	CCAAAGATTG
CAACACGGAG	CAAAGTAAAT	GGGCAATCTG	GCAGGATGGA	GTTTTTCTGG
ACAATCTTAA	AACCCAACGA	TGCGATAAAT	TTCGAGTCCA	ATGGCAATTT
CATCGCCCCT	GAATACGCCT	ATAAGATCGT	GAAAAAGGGG	GACTCTGCAA
TTATGAAGTC	CGAATTAGAG	TATGGCAATT	GCAACACGAA	GTGCCAGACA
CCAATGGGAG	CCATTAATAG	CTCAATGCCC	TCCATAATA	TTCATCCATT
GACCATTGGG	GAGTGCCCAA	AGTACGTGAA	GTCCAACCGC	CTGGTCCTCG
CAACCGGTCT	AAGAAATAGC	CCGCAGAGAG	AATCGCGGAG	GAAGAAACGT
GGCCTGTTTG	GCGCGATTGC	CGGATTCATC	GAGGGAGGCT	GGCAGGGTAT
GGTCGATGGT	TGGTACGGAT	ACCACCATAG	CAACGAACAG	GGGTCCGGCT

ATGCAGCAGA TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAAC
 AAGGTTAATA GTATTATAGA TAAAATGAAC ACGCAATTCG AGGCCGTGG
 GAGGGAGTTT AACAACTCTGG AACGCCGGAT CGAAAATCTG AATAAGAAAA
 TGGGAAGACGG CTTCCCTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA
 CTCATGGAGA ACGAGAGGAC CCTGGATTC CACGATAGCA ACGTGAAGAA
 1401 CCTTTACGAC AAGGTGAGAC TTCAGCTCCG AGACAACGCC
 AAGGAGCTGGGGAATGGATG CTTGAGTTT TACCACAAAT GTGACAATGA
 GTGCATGGAAAGTATACGCA ACGGGACCTA CAATTACCCT CAGTATAGCG
 AAGAGGCTCGGCTCAAACGC GAAGAGATAA GCGGGGTGAA ATTGGAATCA
 ATCGGAACATATCAAATCCT GTCCATCTAT TCCACCGTCG CCTCTTCGCT
 GGCCCTCGCT
 ATCATGATGG CTGGTCTGTC CCTATGGATG TGTTCCAATG
 GAAGCCTTCAGTGCCGTATT TGTATATGAg c
GGCCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCC
TT
 CTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCCAC
 TGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
 TGGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCT
 CTATGGCTTCTGAGGCGGAAAGAACC**AAAGCT**taacatcatc aataatatac
 cttattttgg attgaagcca atatgataat gaggggggtgg
 agtttgtgac gtggcgcggg gcggtgggaac ggggcggggtg acgtagtagt
 gtggcggaag
 tgtgatggtg caagtgtggc ggaacacatg taagcgacgg atgtggcaaa
 agtgacgttt
 ttggtgtgcg cgggtgtaca caggaagtga caattttcgc gcggttttag
 gcggatggtg
 tagtaaattt gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa
 ctgaataaga
 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg taatactgta
 atagtaatca
 attacggggt cattagttca tagcccatat atggagttcc gcggtacata
 acttacggta
 aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat
 aatgacgtat
 gttcccatag taacgccaat aggactttc cattgacgtc aatgggtgga
 gtatttacgg
 taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc
 ccctattgac
 gtcaatgacg gtaaattggc cgctggcat tatgcccagt acatgacctt
 atgggacttt
 cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat
 gcggttttgg
 cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag
 tctccacccc
 attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc
 aaaatgtcgt
 aacaactcgg cccattgac gcaaatgggc ggtagggcgtg tacggtggga
 ggtctatata

agcagagctg gtttagtgaa cgcgcagatc
 cgctag**AGATCTGG**gaaacgatatgggctgaatac**ggatccc**gtattcagcccatatcgtttc
TCT AGA AAT AAA ATA TCT TTA TTT TCA TTA CAT CTG TGT GTT GGT
 TTT TTG TGT GGC
GG CCGCctcagagcct **AAGCTT**ctag ataagatatac cgatccaccg gatctagata
 actgatcata
 atcagccata ccacatttgt agaggtttta cttgctttaa aaaacctccc
 acacctcccc
 ctgaacctga aacataaaat gaatgcaatt gttgttgta acttgtttat
 tgcagcttat
 aatggttaca aataaagcaa tagcatcaca aatttcacaa ataaagcatt
 tttttcactg
 cattctagtt gtggtttgtc caaactcatc aatgtatctt aacgcggatc
 tgggcgtggt
 taagggtggg aaagaatata taagggtggg gtcttatgta gtttt**gtatc**
tgttttgcag
cagcgcgcgc cgccatgagc accaactcgt ttgatggaag cattgtg
AGCTTgtcgactcg aagatctggg cgtggttaag
 ggtgggaaag aatatataag gtgggggtct tatgtagttt tgtatctggt
 ttgcagcagc
 cgccgcgcgc atgagcacca actcgtttga tgggaagcatt gtgagctcat
 atttgacaac
 gcgcatgccc ccattgggccc ggggtgcgtca gaatgtgatg ggctccagca
 ttgatggctg
 ccccgctctg cccgcaaact ctactacctt gacctacgag accgtgtctg
 gaacgcccgtt
 ggagactgca gcctccgccc cgcttcagc cgctgcagcc accgcccgcg
 ggattgtgac
 tgactttget ttctgagcc cgcttgcaag cagtgcagct tcccgttcat
 ccgcccgcga
 tgacaagttg acggctcttt tggcacaatt ggattctttg acccgggaac
 ttaatgtcgt
 ttctcagcag ctgttggatc tgcgccagca ggtttctgcc ctgaaggctt
 cctccccctc
 caatgcggtt taaaacataa ataaaaaacc agactctggt tggatttggg
 tcaagcaagt
 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag gcccgggacc
 agcggctctg
 gtcgttgagg gtccctgtgta tttttccag gacgtggtaa aggtgactct
 ggatgttcag
 atacatgggc ataagcccgt ctctgggggtg gaggtagcac cactgcagag
 cttcatgctg
 cggggtggtg ttgtagatga tccagtcgta gcaggagcgc tgggcgtggt
 gcctaaaaat
 gtctttcagt agcaagctga ttgccagggg caggcccttg gtgtaagtgt
 ttacaaagcg
 gttaagctgg gatgggtgca tacgtgggga tatgagatgc atcttggact
 gtatttttag
 gttggctatg ttcccagcca tatccctccg gggattcatg ttgtgcagaa
 ccaccagcac

agtgtatccg gtgcacttgg gaaatttgtc atgtagctta gaaggaaatg
 cgtggaagaa
 cttggagacg cccttgtgac ctccaagatt ttccatgcat tcgtccataa
 tgatggcaat
 gggcccacgg gcggcggcct gggcgaagat atttctggga tcaactaacgt
 catagtgtg
 ttccaggatg agatcgatc aggccatttt taciaagcgc gggcggaggg
 tgccagactg
 cggataatg gttccatccg gcccaggggc gtagttacc tcacagattt
 gcatttcca
 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg gcgatgaaga
 aaacggttc
 cgggtaggg gagatcagct ggaagaaag caggttcctg agcagctgcg
 acttaccgca
 gccggtgggc ccgtaaata cacctattac cgggtgcaac tggtagttaa
 gagagct
g cagctgcctg catccctgag caggggggccc
 acttcgttaa gcatgtccct gactcgcag tttccctga ccaatccgc
 cagaaggcgc
 tcgccgccca gcgatagcag ttcttgcaag gaagcaaagt tttcaacgg
 tttgagaccg
 tccgccgtag gcatgctttt gagcgtttga ccaagcagtt ccaggcggtc
 ccacagctcg
 gtcacctgct ctacggcatc tcgatccagc atatctctc gtttcgcggg
 ttggggcggc
 tttcgctgta cggcagtagt cggtgctcgt ccagacgggc cagggtcag
 tctttccaag
 ggccagggg cctcgtcagc gtagtctggg tcacggtgaa ggggtgcgct
 ccgggctgcg
 cgctggccag ggtgcgcttg aggctggctc tgctgggtgct gaagcgcctgc
 cggctctcgc
 cctgcgcgtc ggccaggtag catttgacca tgggtgcata gtccagcccc
 tccgcggcgt
 ggcccttggc gcgcagcttg cccttggagg aggcgcgcga cgaggggcag
 tgcagacttt
 tgagggcgta gagcttgggc gcgagaaata ccgattccgg ggagtaggca
 tccgcgcgc
 aggccccgca gacggctctg cattccaaga gccagggtgag ctctggccgt
 tcgggggtcaa
 aaaccagggt tccccatgc tttttgatgc gtttcttacc tctggtttcc
 atgagccggt
 gtccacgctc ggtgacgaaa aggctgtccg tgtccccgta tacagacttg
 agaggcctgt
 cctcgagcgg tgttccgcgg tctcctcgt atagaaactc ggaccactct
 gagacaaagg
 ctcgcgcca gccagcacg aaggaggcta agtgggaggg gtagcggctc
 ttgtccacta
 gggggtccac tcgctccagg gtgtgaagac acatgtgcc ctcttcggca
 tcaaggaagg
 tgattggttt gtaggtgtag gccacgtgac cgggtgttcc tgaagggggg
 ctataaaagg

ggggtgggggc gcgttcgtcc tcaactctctt ccgcatoctt gtctgcgagg
 gccagctggt
 ggggtgagta ctccctctga aaagcgggca tgacttctgc gctaagattg
 tcagtttcca
 6301 aaaacgagga ggatttgata ttcacctggc ccgcggtgat
 gcctttgagg gtggccgcat
 6361 ccatctggtc agaaaagaca atctttttgt tgtcaagctt
 ggtggcaaac gacccgtaga
 6421 gggcgttga cagcaacttg gcgatggagc gcagggtttg
 gtttttgcg cgatcggcgc
 6481 gtccttggc cgcgatgtt agctgcacgt attcgcgcgc
 aacgcaccgc cattcgggaa
 6541 agacggtggc gcgctcgtcg ggcaccaggt gcacgcgcca
 accgcggttg tgcagggtga
 6601 caaggtcaac gctgggtggc acctctccgc gtaggcgctc
 gttggtccag cagaggcggc
 6661 cgcccttgcg cgagcagaat ggcggtaggg ggtctagctg
 cgtctcgtcc ggggggtctg
 6721 cgtccacggc aaagaccccg ggcagcaggc gcgcgtcgaa
 gtagtctatc ttgcatcctt
 6781 gcaagtctag cgctcgtgc catgcgcggg cggcaagcgc
 gcgctcgtat gggttgagtg
 6841 ggggaccca tggcatgggg tgggtgagcg cggaggcgta
 catgccgcaa atgtcgtaaa
 6901 cgtagagggg ctctctgagt attccaagat atgtagggta
 gcatcttcca ccgcggatgc
 6961 tggcgcgcac gtaatcgtat agttcgtgcg agggagcgag
 gaggtcggga ccgaggttgc
 7021 tacgggcggg ctgctctgct cggaaagacta tctgcctgaa
 gatggcatgt gagttggatg
 7081 atatggttgg acgctggaag acggtgaagc tggcgtctgt
 gagacctacc gcgtcacgca
 7141 cgaaggaggc gtaggagtcg cgcagcttgt tgaccagctc
 ggcggtgacc tgcacgtcta
 7201 gggcgcagta gtccagggtt tccttgatga tgtcatactt
 atcctgtccc ttttttttcc
 7261 acagctcgcg gttgaggaca aactcttcgc ggtctttcca
 gtactcttgg atcggaaacc
 7321 cgtcggcctc cgaacggtaa gagcctagca tgtagaactg
 gttgacggcc tggtaggcgc
 7381 agcatcctt ttctacgggt agcgcgtatg cctgcgcggc
 cttccggagc gaggtgtggg
 7441 tgagcgcaaa ggtgtccctg accatgactt tgagggtactg
 gtatttgaag tcagtgtcgt
 7501 cgcacccgcc ctgctcccag agcaaaaagt ccgtgcgctt
 tttggaacgc ggatttgga
 7561 gggcgaaggt gacatcgttg aagagtatct ttcccgcgcg
 aggcataaag ttgcgtgtga
 7621 tgcggaaggg tcccggcacc tcggaacggc tgtaattac
 ctgggcggcg agcacgatct

7681 cgtcaaagcc gttgatgttg tggcccacaa tgtaaagttc
 caagaagcgc gggatgccct
 7741 tgatggaagg caatttttta agttcctcgt aggtgagctc
 ttcaggggag ctgagcccgt
 7801 gctctgaaag ggcccagtct gcaagatgag ggttggaaagc
 gacgaatgag ctccacaggt
 7861 cacgggccat tagcatttgc aggtggtcgc gaaaggtcct
 aaactggcga cctatggcca
 7921 ttttttctgg ggtgatgcag tagaaggtaa gcggttcttg
 ttcccagcgg tcccatccaa
 7981 ggttcgcggc taggtctcgc gcggcagtca ctgagaggctc
 atctccgccg aacttcatga
 8041 ccagcatgaa gggcacgagc tgcttcccaa aggccccat
 ccaagtatag gtctctacat
 8101 cgtaggtgac aaagagacgc tcggtgcgag gatgagcagcc
 gatcgggaag aactggatct
 8161 cccgccacca attggaggag tggctattga tgtggtgaaa
 gtagaagtcc ctgagcggg
 8221 ccgaacactc gtgctggcct ttgtaaaaac gtgagcagta
 ctggcagcgg tgcacgggct
 8281 gtacatcctg cacgagggtg acctgacgac cgcgcacaag
 gaagcagagt gggaatttga
 8341 gccctcgcgc tggcgggttt ggctggtggt cttctacttc
 ggctgcttgt ccttgaccgt
 8401 ctggctgctc gaggggagtt acggtggatc ggaccaccac
 gccgcgcgag cccaaagtcc
 8461 agatgtccgc gcgcggcggc cggagcttga tgacaacatc
 gcgcagatgg gagctgtcca
 8521 tggctctggag ctcccgcggc gtcaggtcag gcgggagctc
 ctgaggttt acctgcata
 8581 gacgggtcag ggcgcgggct agatccaggt gatacctaat
 ttccaggggc tggttggtgg
 8641 cggcgtcgat ggcttgcaag aggccgcac cccgcggcgc
 gactacggta ccgcgcggcg
 8701 ggcgggtggc gcgcgggggtg tccttggatg atgcatctaa
 aagcggtgac gcgggcgagc
 8761 ccccggaggt agggggggct ccggaccgc cgggagaggg
 ggcaggggca cgtcggcgc
 8821 gcgcgcgggc aggagctggt gctgcgcgcg taggttgctg
 gcgaacgcga cgacgcggcg
 8881 gttgatctcc tgaatctggc gcctctgcgt gaagacgacg
 ggcccgggtga gcttgagcct
 8941 gaaagagagt tcgacagaat caatttcggt gtcggtgacg
 gcggcctggc gcaaaatctc
 9001 ctgacgtct cctgagttgt cttgataggc gatctcggcc
 atgaactgct cgatctcttc
 9061 ctcttgaga tctccgcgtc cggctcgcct cacggtggcg
 gcgaggtcgt tggaaatgcg
 9121 ggccatgagc tgcgagaagg cgttgaggcc tccctcgttc
 cagacgcggc ttagaccac

9181 gccccctteg geatcgcggg cgcgcatgac cacctgcgcg
 agattgagct ccacgtgccg
 9241 ggcgaagacg gcgtagtttc gcaggcgctg aaagaggtag
 ttgagggtgg tggcggtgtg
 9301 ttctgccacg aagaagtaca taaccacgagc tgcgaacgtg
 gattcgttga tatccccaa
 9361 ggctcaagg cgctccatgg cctcgtagaa gtccacggcg
 aagttgaaaa actgggagtt
 9421 gcgcgccgac acggttaact cctcctccag aagacggatg
 agctcggcga cagtgtcgcg
 9481 cacctcgcgc tcaaaggcta caggggcctc ttctttctct
 tcaatctcct ctccataag
 9541 ggctccccct tctttcttct ctggcgggcg tgggggaggg
 gggacacggc ggcgacgacg
 9601 gcgcaccggg aggcggtcga caaagcgctc gatcatctcc
 ccgcggcgac ggcgcatggt
 9661 ctcggtgacg gcgcggccgt tctcgcgggg gcgcagttgg
 aagacgccgc ccgtcatgtc
 9721 ccggttatgg gttggcgggg ggctgccatg cggcagggat
 acggcgctaa cgatgcatct
 9781 caacaattgt tgtgtaggta ctccgccgcc gagggacctg
 agcgagtccg catcgaccgg
 9841 atcggaaaac ctctcgagaa aggcgtctaa ccagtcacag
 tcgcaaggta ggctgagcac
 9901 cgtggcgggc ggcagcgggc ggcggtcggg gttgtttctg
 gcggaggtgc tgctgatgat
 9961 gtaattaaag taggcggtct tgagacggcg gatggtcgac
 agaagcacca tgtccttggg
 10021 tccggcctgc tgaatgcgca ggcggtcggc catgccccag
 gcttcgtttt gacatcggcg
 10081 caggtctttg tagtagtctt gcatgagcct ttctaccggc
 acttcttctt ctcttctc
 10141 ttgtcctgca tctcttgcac ctatcgctgc ggcggcgggc
 gagtttggcc gtaggtggcg
 10201 ccctcttct cccatgcgtg tgaccccgaa gccctcctc
 ggctgaagca ggcctaggtc
 10261 ggcgacaacg cgctcggcta atatggcctg ctgcacctgc
 gtgagggtag actggaagtc
 10321 atccatgtcc acaaagcggg ggtatgcgcc cgtgttgatg
 gtgtaagtgc agttggccat
 10381 aacggaccag ttaacggtct ggtgaccggg ctgcgagagc
 tcggtgtacc tgagacgcga
 10441 gtaagccctc gagtcaaata cgtagtcggt gcaagtccgc
 accaggtact ggtatcccac
 10501 caaaaagtgc ggcggcggtt ggcggtagag gggccagcgt
 aggggtggccg gggctccggg
 10561 ggcgagatct tccaacataa ggcgatgata tccgtagatg
 tacctggaca tccaggtgat
 10621 gccggcgcg gtggtggagg cgcgcgaaa gtcgcggaacg
 cggttccaga tgttgcgcag

10681 cggcaaaaag tgctccatgg tggggacgct ctggccggtc
 aggcgcgcgc aatcgttgac
 10741 gctctaccgt gcaaaaaggag agcctgtaag cgggcactct
 tccgtggtct ggtggataaa
 10801 ttcgcaaggg tatcatggcg gacgaccggg gttcgagccc
 cgtatccggc cgtccgccgt
 10861 gatccatgcg gttaccgccc gcgtgtcgaa cccaggtgtg
 cgacgtcaga caacggggga
 10921 gtgctccttt tggcttctt ccaggcgcgg cggctgctgc
 gctagctttt ttggccactg
 10981 gccgcgcgca gcgtaagcgg ttaggctgga aagcgaaagc
 attaagtggc tcgctccctg
 11041 tagccggagg gttattttcc aagggttgag tcgcgggacc
 cccggttcga gtctcggacc
 11101 ggccggactg cggcgaacgg gggtttgct ccccgatcag
 caagaccccg cttgcaaatt
 11161 cctccgaaa cagggacgag cccctttttt gcttttcca
 gatgcatccg gtgctgcggc
 11221 agatgccc ccctcctcag cagcggcaag agcaagagca
 gcggcagaca tgcagggcac
 11281 cctcccctcc tctaccgcg tcaggagggg cgacatccgc
 ggttgacgcg gcagcagatg
 11341 gtgattacga acccccggg cgcggggccc ggcactacct
 ggacttgag gagggcgagg
 11401 gcctggcgcg gctaggagcg ccctctctg agcggtaacc
 aagggtgcag ctgaagcgtg
 11461 atacgcgtga ggcgtacgtg ccgcggcaga acctgtttcg
 cgaccgcgag ggagaggagc
 11521 ccgaggagat gcgggatcga aagtccacg cagggcgcga
 gctgcggcat ggcctgaatc
 11581 gcgagcggtt gctgcgcgag gaggactttg agcccgcgc
 gcgaaccggg attagtcccg
 11641 cgcgcgcaca cgtggcggcc gccgacctgg taaccgcata
 cgagcagacg gtgaaccagg
 11701 agattaactt tcaaaaagc tttacaacc acgtgcgtac
 gcttggtggc gcgaggagg
 11761 tggctatagg actgatgcat ctgtgggact ttgtaagcgc
 gctggagcaa aaccxaaata
 11821 gcaagccgct catggcgcag ctgttcctta tagtgcagca
 cagcagggac aacgaggcat
 11881 tcagggatgc gctgctaaac atagtagagc ccgagggccg
 ctggctgctc gatttgataa
 11941 acatcctgca gagcatagtg gtgcaggagc gcagcttgag
 cctggctgac aagggtggccg
 12001 ccatcaacta ttccatgctt agcctgggca agttttacgc
 ccgcaagata taccataccc
 12061 cttacgttcc catagacaag gaggtaaaga tcgagggggtt
 ctacatgcgc atggcgtga
 12121 aggtgcttac cttgagcgc gacctgggcg tttatcgcaa
 cgagcgcac cacaaggccg

12181 tgagcgtgag cggcgggcgc gagctcagcg accgcgagct
 gatgcacagc ctgcaaaggg
 12241 ccctggctgg cacgggcagc ggcgatagag aggccgagtc
 ctactttgac gcgggcgctg
 12301 acctgcgctg ggccccaagc cgacgcgccc tggaggcagc
 tggggccgga cctgggctgg
 12361 cgggtggcacc cgcgcgcgct ggcaacgtcg gcggcgtgga
 ggaatatgac gaggacgatg
 12421 agtacgagcc agaggacggc gagtactaag cggtgatgtt
 tctgatcaga tgatgcaaga
 12481 cgcaacggac ccggcgggtgc gggcggcgct gcagagccag
 ccgtccggcc ttaactccac
 12541 ggacgactgg cgccagggtca tggaccgcat catgtcgctg
 actgcgcgca atcctgacgc
 12601 gttccggcag cagccgcagg ccaaccggct ctccgcaatt
 ctggaagcgg tgggtcccggc
 12661 gcgcgcaaac cccacgcacg agaaggtgct ggcgatcgta
 aacgcgctgg ccgaaaacag
 12721 ggccatccgg cccgacgagg ccggcctggt ctacgacgcg
 ctgcttcagc gcgtggctcg
 12781 ttacaacagc ggcaacgtgc agaccaacct ggaccggctg
 gtgggggatg tgcgcgaggc
 12841 cgtggcgcag cgtgagcgcg cgcagcagca gggcaacctg
 ggctccatgg ttgcactaaa
 12901 cgctttcctg agtacacagc ccgccaacgt gccgcgggga
 caggaggact acaccaactt
 12961 tgtgagcgcga ctgcccgtaa tgggtgactga gacaccgcaa
 agtgaggtgt accagtctgg
 13021 gccagactat tttttccaga ccagtagaca aggcctgcag
 accgtaaacc tgagccaggc
 13081 tttcaaaaac ttgcaggggc tgtgggggggt gcgggctccc
 acaggcgacc gcgcgaccgt
 13141 gtctagcttg ctgacgcca actcgcgcct gttgctgctg
 ctaatagcgc ctttcacgga
 13201 cagtggcagc gtgtcccggg acacatacct aggtcacttg
 ctgacactgt accgcgaggc
 13261 cataggtcag gcgcatgtgg acgagcatac tttccaggag
 attacaagtg tcagccgcgc
 13321 gctggggcag gaggacacgg gcagcctgga ggcaacccta
 aactacctgc tgaccaaccg
 13381 gcggcagaag atcccctcgt tgcacagttt aaacagcgag
 gaggagcgcga ttttgcgcta
 13441 cgtgcagcag agcgtgagcc ttaacctgat gcgcgacggg
 gtaacgcca gcgtggcgct
 13501 ggacatgacc gcgcgcaaca tggaaccggg catgtatgcc
 tcaaaccggc cgtttatcaa
 13561 ccgcctaattg gactacttgc atcgcgcggc cgccgtgaac
 cccgagtatt tcaccaatgc
 13621 catcttgaac ccgcaactggc taccgcccc tggtttctac
 accgggggat tcgaggtgcc

13681 cgagggtaac gatggattcc tctgggacga catagacgac
 agcgtgtttt ccccgcaacc
 13741 gcagaccctg ctagagtgc aacagcgcgga gcaggcagag
 gcggcgctgc gaaaggaag
 13801 cttccgcagg ccaagcagct tgtccgatct aggcgctgcg
 gccccgcggt cagatgctag
 13861 tagcccattt ccaagcttga tagggtctct taccagcact
 cgcaccaccc gcccggcct
 13921 gctgggagag gaggagtacc taaacaactc gctgctgcag
 ccgcagcgcg aaaaaaacct
 13981 gcctccggca tttcccaaca acgggataga gagcctagtg
 gacaagatga gtagatgaa
 14041 gacgtacgcg caggagcaca gggacgtgcc aggcccgcg
 ccgcccaccc gtcgtcaaag
 14101 gcacgaccgt cagcggggtc tgggtgtggga ggacgatgac
 tcggcagacg acagcagcgt
 14161 cctggatttg ggaggagtg gcaaccggtt tgcgcacctt
 cgccccaggc tggggagaat
 14221 gttttaaaaa aaaaaaagca tgatgcaaaa taaaaaactc
 accaaggcca tggcaccgag
 14281 cgttggtttt cttgtattcc ccttagtatg cggcgcgcgg
 cgatgtatga ggaaggctct
 14341 cctccctcct acgagagtgt ggtgagcgcg gcgccagtgg
 cggcggcgct gggttctccc
 14401 ttcgatgctc cctggacc cccgtttgct cctccgcggt
 acctgcggcc taccgggggg
 14461 agaaacagca tccgttactc tgagttggca cccctattcg
 acaccaccg tgtgtacctg
 14521 gtggacaaca agtcaacgga tgtggcatcc ctgaactacc
 agaacgacca cagcaacttt
 14581 ctgaccacgg tcattcaaaa caatgactac agcccggggg
 aggcaagcac acagaccatc
 14641 aatcttgacg accggtcgca ctggggcggc gacctgaaaa
 ccatcctgca taccaacatg
 14701 ccaaagtga acgagttcat gttaccaat aagtttaagg
 cgcgggtgat ggtgtcgcgc
 14761 ttgcctacta aggacaatca ggtggagctg aaatacgagt
 ggggtggagtt cacgctgccc
 14821 gagggcaact actccgagac catgaccata gaccttatga
 acaacgcgat cgtggagcac
 14881 tacttgaaag tgggcagaca gaacgggggt ctggaaagcg
 acatcgggggt aaagtttgac
 14941 acccgcaact tcagactggg gtttgacccc gtcactggtc
 ttgtcatgcc tggggtatat
 15001 acaaacgaag ccttccatcc agacatcatt ttgctgccag
 gatgcgggggt ggacttcacc
 15061 cacagccgcc tgagcaactt gttgggcatc cgcaagcggc
 aacccttcca ggagggttt
 15121 aggatcacct acgatgatct ggagggtggg aacattcccg
 cactgttggga tgtggacgcc

15181 taccaggcga gcttgaaaga tgacaccgaa cagggcgggg
 gtggcgcagg cggcagcaac
 15241 agcagtggca gcggcgcgga agagaactcc aacgcggcag
 ccgcggcaat gcagccggtg
 15301 gaggacatga acgatcatgc cattcgcggc gacaccttg
 ccacacgggc tgaggagaag
 15361 cgcgctgagg ccgaagcagc ggccgaagct gccgccccg
 ctgcgcaacc cgaggtcgag
 15421 aagcctcaga agaaaccggt gatcaaacc ctgacagagg
 acagcaagaa acgcagttac
 15481 aacctaataa gcaatgacag caccttcacc cagtaccgca
 gctggtacct tgcatacaac
 15541 tacggcgacc ctcagaccgg aatccgctca tggaccctgc
 tttgcaactcc tgacgtaacc
 15601 tgcggctcgg agcaggtcta ctggtcggtg ccagacatga
 tgcaagacc cgtgacctc
 15661 cgctccacgc gccagatcag caactttccg gtggtgggcg
 ccgagctggt gcccgtcac
 15721 tccaagagct tctacaacga ccaggccgtc tactcccaac
 tcatccgcca gtttacctct
 15781 ctgaccacg tgttcaatcg ctttcccgag aaccagattt
 tggecgccc gccagcccc
 15841 accatcacca ccgtcagtga aaacgttcct gctctcacag
 atcacgggac gctaccgctg
 15901 cgcaacagca tcggaggagt ccagcgagt accattactg
 acgccagacg ccgcacctgc
 15961 ccctacgttt acaaggccct gggcatagtc tcgcccgcg
 tcctatcgag ccgcactttt
 16021 tgagcaagca tgtccatcct tatatcgccc agcaataaca
 caggctgggg cctgcgcttc
 16081 ccaagcaaga tgtttgccg ggccaagaag cgctccgacc
 aacaccagc gcgcgtgcg
 16141 gggcactacc gcgcgcctg gggcgcgcac aaacgcggcc
 gactgggcg caccaccgtc
 16201 gatgacgcca tcgacgcggt ggtggaggag gcgcgcaact
 acacgccac gccgccacca
 16261 gtgtccacag tggacgcggc cattcagacc gtggtgcgcg
 gagcccggcg ctatgctaaa
 16321 atgaagagac ggcggaggcg cgtagcacgt cgccaccgcc
 gccgaccg cactgccgc
 16381 caacgcgcg gcgggcctt gcttaaccgc gcacgtcgca
 ccggccgacg ggcggccatg
 16441 cgggcccgtc gaaggctggc gcgggtatt gtcactgtgc
 ccccaggtc caggcgacga
 16501 gcggccgcg cagcagccgc gccattagt gctatgactc
 agggtcgacg gggcaacgtg
 16561 tattgggtgc gcgactcggg tagcggcctg cgcgtgcccg
 tgcgcacccg cccccgcg
 16621 aactagattg caagaaaaa ctacttagac tcgtactggt
 gtatgtatcc agcggcggcg

16681 gcgcgcaacg aagctatgtc caagcgcaaa atcaaagaag
 agatgctcca ggtcatcgcg
 16741 ccggagatct atggccccc gaagaaggaa gagcaggatt
 acaagccccg aaagctaaag
 16801 cgggtcaaaa agaaaaagaa agatgatgat gatgaacttg
 acgacgaggt ggaactgctg
 16861 cacgctaccg cgcccaggcg acgggtacag tggaaaggtc
 gacgcgtaaa acgtgttttg
 16921 cgaccggca ccaccgtagt ctttacgccc ggtgagcgct
 ccaccgcac ctacaagcgc
 16981 gtgtatgatg aggtgtacgg cgacgaggac ctgcttgagc
 aggccaacga ggcctcggg
 17041 gagtttgctt acggaaagcg gcataaggac atgctggcgt
 tgccgctgga cgagggaac
 17101 ccaacaccta gcctaaagcc cgtaaacctg cagcagggtc
 tgcccgcgct tgcaccgtcc
 17161 gaagaaaagc gcggcctaaa gcgcgagtct ggtgacttgg
 caccaccgt gcagctgatg
 17221 gtaccaagc gccagcgact ggaagatgct ttggaaaaa
 tgaccgtgga acctgggctg
 17281 gagcccagg tccgcgtgcg gccaatcaag caggtggcgc
 cgggactggg cgtgcagacc
 17341 gtggacgttc agataccac taccagtagc accagtattg
 ccaccgccac agagggcatg
 17401 gagacacaaa cgtccccggt tgccctcagcg gtggcggatg
 ccgcggtgca ggcggtcgt
 17461 gcggccgct ccaagacctc tacggagggtg caaacggacc
 cgtggatggt tcgcgttca
 17521 gccccccggc gcccgcgcgg ttcgaggaag tacggcgccg
 ccagcgcgct actgcccga
 17581 tatgcctac atccttccat tgcgcctacc cccggctatc
 gtggctacac ctaccgcccc
 17641 agaagacgag caactaccg acgccgaacc accactggaa
 ccgcccggc cgctcgcgt
 17701 cgccagcccg tgctggcccc gatttccgtg cgcagggtg
 ctcgcgaagg aggcaggacc
 17761 ctggtgctgc caacagcgcg ctaccacccc agcatcgttt
 aaaagccggt ctttgtggtt
 17821 cttgcagata tggccctcac ctgcccctc cgtttcccgg
 tgcccgggatt ccgaggaaga
 17881 atgcaccgta ggaggggcat ggccggccac ggccctgaegg
 gcggcatgcg tcgtgcgcac
 17941 caccggcggc ggcgcgcgct gcaccgtcgc atgcgcggcg
 gtatcctgcc ctccttatt
 18001 cactgatcg ccgcgcgat tggcgccgtg cccggaattg
 catccgtggc cttgcaggcg
 18061 cagagacact gattaaaaac aagttgcatg tggaaaaatc
 aaaataaaaa gtctggactc
 18121 tcagctcgc ttggtcctgt aactattttg tagaatggaa
 gacatcaact ttgcgtctct

18181 ggccccgcga cacggctcgc gcccgttcat gggaaactgg
 caagatatcg gcaccagcaa
 18241 tatgagcggg ggcgccttca gctggggctc gctgtggagc
 ggcattaanaa atttcggttc
 18301 caccgttaag aactatggca gcaaggcctg gaacagcagc
 acaggccaga tgctgagggg
 18361 taagttgaaa gagcaaaatt tccaacaaaa ggtggtagat
 ggcttggcct ctggcattag
 18421 cgggggtggg gacctggcca accaggcagt gcaaaaataag
 attaacagta agcttgatcc
 18481 ccgccctccc gtagaggagc ctccaccggc cgtggagaca
 gtgtctccag aggggctggg
 18541 cgaaaagcgt ccgcgccccg acagggaga aactctggg
 acgcaaatag acgagcctcc
 18601 ctcgtagag gaggcactaa agcaaggcct gccaccacc
 cgtcccatcg cgcccatggc
 18661 taccggagtg ctgggcccagc acacaccctg aacgctggac
 ctgcctcccc ccgcccagac
 18721 ccagcagaaa cctgtgctgc caggcccagc cgccgttgtt
 gtaaccctgc ctagccgcgc
 18781 gtccctgcgc cgcgcccgcca gcggtccgcg atcgttgccg
 cccgtagcca gtggcaactg
 18841 gcaaagcaca ctgaacagca tcgtgggtct gggggtgcaa
 tccctgaagc gccgacgatg
 18901 cttctgaata gctaacgtgt cgtatgtgtg tcatgtatgc
 gtccatgtcg ccgccagagg
 18961 agctgctgag ccgcccgcgc cccgctttcc aagatggcta
 ccccttcgat gatgccgcag
 19021 tggctttaca tgcacatctc gggccaggac gcctcggagt
 acctgagccc cgggctggg
 19081 cagtttgccc gcgccaccga gacgtacttc agcctgaata
 acaagtttag aaaccccacg
 19141 gtggcgccta cgcacgacgt gaccacagac cggctcccagc
 gtttgacgct gcggttcac
 19201 cctgtggacc gtgaggatac tgcgtactcg tacaaggcgc
 ggttcaccct agctgtggg
 19261 gataaccgtg tgctggacat ggcttccacg tactttgaca
 tccgcccgcgt gctggacagg
 19321 ggccctactt ttaagcccta ctctggcact gcctacaacg
 ccctggctcc caaggggtgc
 19381 ccaatcctt gcgaatggga tgaagctgct actgctcttg
 aaataaacct agaagaagag
 19441 gacgatgaca acgaagacga agtagacgag caagctgagc
 agcaaaaaac tcacgtattt
 19501 gggcaggcgc cttattctgg tataaatatt acaaaggagg
 gtattcaaat aggtgtcgaa
 19561 ggtaaacac ctaaataatgc cgataaaaaca tttcaacctg
 aacctcaaat aggagaatct
 19621 cagtggtagc aaactgaaat taatcatgca gctgggagag
 tccttaaaaa gactaccca

19681 atgaaacccat gttacgggttc atatgcaaaa cccacaaatg
 aaaatggagg gcaaggcatt
 19741 cttgtaaagc acaaaaatgg aaagctagaa agtcaagtgg
 aatgcaatt tttctcaact
 19801 actgaggcga ccgcaggcaa tggtgataac ttgactccta
 aagtgggtatt gtacagttaa
 19861 gatgtagata tagaaacccc agacactcat atttcttaca
 tgcccactat taaggaaggt
 19921 aactcacgag aactaatggg ccaacaatct atgcccaca
 ggcctaatta cattgctttt
 19981 agggacaatt ttattgggtct aatgtattac aacagcacgg
 gtaatatggg tgttctggcg
 20041 ggccaagcat cgcagttgaa tgctgttgta gatttgcaag
 acagaaacac agagctttca
 20101 taccagcttt tgcttgatc cattggtgat agaaccaggt
 acttttctat gtggaatcag
 20161 gctggtgaca gctatgatcc agatgtaga attattgaaa
 atcatggaac tgaagatgaa
 20221 cttccaaatt actgctttcc actgggaggt gtgattaata
 cagagactct taccaaggta
 20281 aaacctaaaa caggtcagga aatggatgg gaaaaagatg
 ctacagaatt ttcagataaa
 20341 aatgaaataa gagttggaaa taattttgcc atggaaatca
 atctaaatgc caacctgtgg
 20401 agaaatttcc tgtaactcaa catagcgctg tatttgcccg
 acaagctaaa gtacagtcct
 20461 tccaacgtaa aaatttctga taacccaaac acctacgact
 acatgaacaa gcgagtgggtg
 20521 gctcccgggt tagtggactg ctacattaac cttggagcac
 gctggtcctt tgactatatg
 20581 gacaacgtca acccatttaa ccaccaccgc aatgctggcc
 tgcgctaccg ctcaatgttg
 20641 ctgggcaatg gtcgctatgt gcccttccac atccaggtgc
 ctacagaagtt ctttgccatt
 20701 aaaaacctcc ttctcctgcc gggctcatac acctacgagt
 ggaacttcag gaaggatggt
 20761 aacatgggtc tgcagagctc cctaggaat gacctaaagg
 ttgacggagc cagcattaag
 20821 tttgatagca tttgccttta cgccacctc tccccatgg
 cccacaacac cgctccacg
 20881 cttgaggcca tgcttagaaa cgacaccaac gaccagtcct
 ttaacgacta tctctccgcc
 20941 gccaacatgc tctaccctat acccgccaac gctaccaacg
 tgcccatac catcccctcc
 21001 cgcaactggg cggctttccg cggctgggccc ttcacgcgcc
 ttaagactaa ggaaaccca
 21061 tcaactgggt cgggctacga cccttattac acctactctg
 gctctatacc ctacctagat
 21121 ggaacctttt acctcaacca cacctttaag aagggtggcca
 ttacctttga ctcttctgtc

21181 agctggcctg gcaatgaccg cctgcttacc cccaacgagt
 ttgaaattaa gcgctcagtt
 21241 gacggggagg gttacaacgt tgcccagtggt aacatgacca
 aagactgggt cctggtacaa
 21301 atgctagcta actacaacat tggctaccag ggcttctata
 tcccagagag ctacaaggac
 21361 cgcattgtact ccttctttag aaacttccag cccatgagcc
 gtcaggtggg gcatgatact
 21421 aaatacaagg actaccaaca ggtgggcatc ctacaccaac
 acaacaactc tggatttggt
 21481 ggctaccttg cccccaccat gcgcgaagga caggcctacc
 ctgctaactt cccctatccg
 21541 cttataggca agaccgcagt tgacagcatt acccagaaaa
 agtttctttg ccatcgcacc
 21601 ctttggcgca tcccattctc cagtaacttt atgtccatgg
 gcgcactcac agacctgggc
 21661 caaaccttc tctacgcaa ctccgcccac gcgctagaca
 tgacttttga ggtggatccc
 21721 atggacgagc ccacccttct ttatgttttg tttgaagtct
 ttgacgtggg ccgtgtgcac
 21781 cggccgcacc gcggcgctcat cgaaaccgtg tacctgcgca
 cgcccttctc ggccggcaac
 21841 gccacaacat aaagaagcaa gcaacatcaa caacagctgc
 cgccatgggc tccagtgagc
 21901 aggaactgaa agccattgtc aaagatcttg gttgtgggcc
 atattttttg ggcacctatg
 21961 acaagcgctt tccaggcttt gtttctccac acaagctcgc
 ctgcgccata gtcaatacgg
 22021 ccggtcgcga gactgggggc gtacactgga tggcctttgc
 ctggaaccgg cactcaaaaa
 22081 catgctacct ctttgagccc tttggctttt ctgaccagcg
 actcaagcag gtttaccagt
 22141 ttgagtaega gtcactcctg cgccgtagcg ccattgcttc
 ttcccccgac cgctgtataa
 22201 cgctggaaaa gtccacccaa agcgtacagg ggcccaactc
 ggccgcctgt ggactattct
 22261 gctgcatggt tctccacgcc tttgccaact ggcccaaac
 tcccatggat cacaacccca
 22321 ccatgaacct tattaccggg gtacccaact ccatgctcaa
 cagtccccag gtacagccca
 22381 cctgcgctcg caaccaggaa cagctctaca gcttctgga
 gcgccactcg cctacttcc
 22441 gcagccacag tgcgcagatt aggagcgcca cttctttttg
 tcacttgaaa aacatgtaa
 22501 aataatgtac tagagacact ttcaataaag gcaaatgctt
 ttatttgtag actctcgggt
 22561 gattatttac ccccacctt gccgtctgcg ccgtttaaaa
 atcaaagggg ttctgcccgc
 22621 catcgctatg cgccactggc agggacacgt tgcgatactg
 gtgttttagtg ctccacttaa

22681 actcaggcac aaccatccgc ggcagctcgg tgaagttttc
 actccacagg ctgcgcacca
 22741 tcaccaacgc gtttagcagg tcgggcgccg atatcttgaa
 gtgcgagttg gggcctccgc
 22801 cctgocgcgc cgagttgcga tacacagggt tgcagcactg
 gaacactatc agcgcgggt
 22861 ggtgcacgct ggccagcacg ctcttgctcg agatcagatc
 cgcgtccagg tctcgcgcg
 22921 tgctcagggc gaacggagtc aactttggta gctgccttc
 caaaaagggc gcgtgccag
 22981 gctttgagtt gcactgcac cgtagtggca tcaaaagggt
 accgtgcccg gtctgggcg
 23041 taggatacag cgctgcata aaagccttga tctgcttaaa
 agccacctga gcctttgcg
 23101 cttcagagaa gaacatgccg caagacttgc cggaaaactg
 attggccgga caggccgcg
 23161 cgtgcacgca gcacctgcg tcggtggtgg agatctgcac
 cacatttcgg ccccaccggt
 23221 tcttcacgat cttggccttg ctagactgct ccttcagcgc
 gcgctgcccg ttttcgctcg
 23281 tcacatccat ttcaatcag tgctccttat ttatcataat
 gcttccgtgt agacacttaa
 23341 gctcgccttc gatctcagcg cagcggtgca gccacaacgc
 gcagcccgtg ggctcgtgat
 23401 gcttgtaggt cacctctgca aacgactgca ggtacgcctg
 caggaatcgc cccatcatcg
 23461 tcacaaagggt cttggtgctg gtgaaggtea gctgcaacce
 gcggtgctcc tcgttcagcc
 23521 aggtcttgca tacggccgcc agagcttcca cttggtcagg
 cagtagtttg aagtccgct
 23581 ttagatcgtt atccacgtgg tacttgteca tcagcgcgcg
 cgcagcctcc atgcccttct
 23641 cccacgcaga cacgatcggc aactcagcg ggttcatcac
 cgtaatttca ctttcgctt
 23701 cgetgggctc ttcctcttcc tcttgctcc gcataccacg
 cgccactggg tcgtcttcat
 23761 tcagccgccg cactgtgcgc ttacctcctt tgccatgctt
 gattagcacc ggtgggtg
 23821 tgaaaaccac catttgtagc gccacatctt ctctttcttc
 ctgctgtcc acgattacct
 23881 ctggtgatgg cgggcgctcg ggcttgggag aagggcgctt
 ctttttcttc ttgggcgcaa
 23941 tggccaaatc cgcgcgcgag gtcgatggcc gcgggctggg
 tgtgcgcggc accagcgcg
 24001 cttgtgatga gtcttctcgc tctcggact cgatacgcgcg
 cctcctccgc ttttttggg
 24061 gcgcccgggg aggcggcggc gacggggacg gggacgacac
 gtctccatg gttgggggac
 24121 gtcgcgcgcg accgcgtccg cgctcggggg tggtttcgcg
 ctgctcctct tcccgaactgg

24181 ccatttcctt ctctatagg cagaaaaaga tcatggagtc
 agtcgagaag aaggacagcc
 24241 taaccgccc ctctgagttc gccaccaccg cctccaccga
 tgccgccaac ggcctacca
 24301 ccttcccgt cgaggcacc cgccttgagg aggaggaagt
 gattatcgag caggaccag
 24361 gttttgtaag cgaagacgac gaggaccgct cagtaccaac
 agaggataaa aagcaagacc
 24421 aggacaacgc agaggcaaac gaggaacaag tcgggcgggg
 ggacgaaagg catggcgact
 24481 acctagatgt gggagacgac gtgctgttga agcatctgca
 gcgccagtgc gccattatct
 24541 gcgacgcgtt gcaagagcgc agcgatgtgc ccctcgccat
 agcggatgtc agccttgcc
 24601 acgaacgcca cctattctca ccgcgctac ccccaaacg
 ccaagaaaac ggcacatgca
 24661 agcccaacc ggcctcaac ttctaccccg tatttgccgt
 gccagaggtg cttgccacct
 24721 atcacatctt tttccaaaac tgcaagatac ccctatcctg
 ccgtgccaac cgcagccgag
 24781 cggacaagca gctggccttg cggcagggcg ctgtcatacc
 tgatatcgcc tcgctcaacg
 24841 aagtgccaaa aatctttgag ggtcttgac gcgacgagaa
 gcgcgcgga aacgctctgc
 24901 aacaggaaa cagcgaat gaaagtcact ctggagtgtt
 ggtggaactc gagggtgaca
 24961 acgcgccct agccgtacta aaacgcagca tcgaggtcac
 ccactttgcc taccggcac
 25021 ttaacctacc cccaaggtc atgagcacag tcatgagtga
 gctgatcgtg cgccgtgca
 25081 agcccctgga gagggatgca aatttgcaag aacaacaga
 ggagggccta cccgagttg
 25141 gcgacgagca gctagcgcg tggcttcaa cgcgagagcc
 tgccgacttg gaggagcga
 25201 gaaaactaat gatggccgca gtgctcgta ccgtggagct
 tgagtgcag cagcgttct
 25261 ttgctgacc ggagatgcag cgcaagctag aggaaacatt
 gactacacc tttcgacag
 25321 gctacgtacg ccaggcctgc aagatctcca acgtggagct
 ctgcaacctg gtctctacc
 25381 ttggaatttt gcacgaaaac cgccttgggc aaaacgtgct
 tcattccag ctcaagggcg
 25441 aggcgcccgc gactacgtc cgcgactgca tttacttatt
 tctatgctac acctggcaga
 25501 cggccatggg cgtttggcag cagtgcttg aggagtgcaa
 cctcaaggag ctgcagaaac
 25561 tgctaaagca aaacttgaag gacctatgga cggccttcaa
 cgagcgtcc gtggccgca
 25621 acctggcgga catcattttc cccgaacgcc tgcttaaac
 cctgcaacag ggtctgccc

25681 acttcaccag tcaaagcatg ttgcagaact ttaggaactt
 tatectagag cgctcaggaa
 25741 tcttgcccgc cacctgctgt gcacttccta gcgactttgt
 gccattaag taccgcgaat
 25801 gccctccgcc gctttggggc cactgctacc ttctgcagct
 agccaactac cttgcctacc
 25861 actctgacat aatggaagac gtgagcgggt acggtctact
 ggagtgtcac tgtcgtgca
 25921 acctatgcac ccgcaccgc tccttgggtt gcaattcgca
 gctgcttaac gaaagtcaaa
 25981 ttatecgtac ctttgagctg cagggtcctt cgcctgacga
 aaagtccgcg gctccggggg
 26041 tgaaactcac tccgggggctg tggacgtcgg cttacctcg
 caaatgtgta cctgaggact
 26101 accacgcca cgagattagg ttctacgaag accaatcccg
 cccgccaaat gcgagctta
 26161 ccgctgctg cattaccag ggccacattc ttggccaatt
 gcaagccatc acaaagccc
 26221 gccaaagatt tctgctacga aaggacggg gggtttactt
 ggaccccag tccggcgagg
 26281 agctcaacc aatcccccg ccgcccagc cctatcagca
 gcagccgcg gcccttgctt
 26341 ccaggatgg cacccaaaa gaagctgcag ctgccgcgc
 caccacgga cgaggaggaa
 26401 tactgggaca gtcaggcaga ggaggttttg gacgaggagg
 aggaggacat gatggaagac
 26461 tgggagagcc tagacgagga agcttccgag gtcgaagagg
 tgtcagacga aacaccgtca
 26521 ccctcggctg cattcccctc gccggcgccc cagaaatcgg
 caaccggttc cagcatggct
 26581 acaacctccg ctctcaggc gccgcccgca ctgcccgttc
 gccgaccaa ccgtagatg
 26641 gacaccactg gaaccagggc cggtaagtcc aagcagccgc
 cgccgttagc ccaagagcaa
 26701 caacagcgc aaggctaccg ctcatggcgc gggcacaaga
 acgcatagt tgcttgcttg
 26761 caagactgtg ggggcaacat ctcttcgcc cgccgctttc
 ttcttacca tcacggcgtg
 26821 gccttcccc gtaacatcct gcattactac cgtcatctct
 acagcccata ctgcaccgpc
 26881 ggcagcggca gcggcagcaa cagcagcggc cacacagaag
 caaaggcgac cggatagcaa
 26941 gactctgaca aagcccaga aatccacagc ggcggcagca
 gcaggaggag gagcgtgca
 27001 tctggcgccc aacgaaccg tatcgaccg cgagcttaga
 aacaggattt tcccactct
 27061 gtatgctata tttcaacaga gcaggggcca agaacaagag
 ctgaaaataa aaaacaggtc
 27121 tctgcgatcc ctaccgcga gctgcctgta tcacaaaagc
 gaagatcagc ttcggcgcac

27181 gctggaagac gcgaggctc tcttcagtaa atactgcgcg
 ctgactctta aggactagtt
 27241 tcgogccctt tctcaaattt aagcgcgaaa actacgtcat
 ctccagcggc cacaccggc
 27301 gccagcacct gtcgtcagcg ccattatgag caaggaaatt
 cccacgcctt acatgtggag
 27361 ttaccagcca caaatgggac ttgcggtgag agctgcccaa
 gactactcaa cccgaataaa
 27421 ctacatgagc gcgggacccc acatgatatc cgggtcaac
 ggaatccgcg cccaccgaaa
 27481 ccgaattctc ttggaacagg cggctattac caccacacct
 cgtaataacc ttaatccccg
 27541 tagttggccc gctgccctgg tgtaccagga aagtcccgt
 cccaccactg tggacttcc
 27601 cagagacgcc caggccgaag ttcagatgac taactcaggg
 gcgcagcttg cgggcccgtt
 27661 tcgtcacagg gtgcggtcgc cggggcaggg tataactcac
 ctgacaatca gagggcgagg
 27721 tattcagctc aacgacgagt cggtgagctc ctgcttggg
 ctccgtccgg acgggacatt
 27781 tcagatcggc ggcgccggcc gtccttcatt cacgcctcgt
 caggcaatcc taactctgca
 27841 gacctcgtcc tctgagccgc gctctggagg cattggaact
 ctgcaattta ttgaggagtt
 27901 tgtgccatcg gtctacttta accccttctc gggacctccc
 ggccactatc cggatcaatt
 27961 tattcctaac ttgacgcgg taaaggactc ggccgacggc
 tacgactgaa tgtaagtgg
 28021 agaggcagag caactgcgcc tgaaacacct ggtccactgt
 cgccgccaca agtgctttgc
 28081 ccgcgactcc ggtgagtttt gctactttga attgcccag
 gatcatatcg agggcccggc
 28141 gcacggcgtc cggcttaccg cccagggaga gcttgcccgt
 agcctgattc gggagtttac
 28201 ccagcgcccc ctgctagttg agcgggacag gggacctgt
 gttctcactg tgatttgcaa
 28261 ctgtcctaac cttggattac atcaagatcc tctagttata
 actagagtac cgggggatct
 28321 tattcccttt aactaataaa aaaaaataat aaagcatcac
 ttacttaaaa tcagttagca
 28381 aatttctgtc cagtttattc agcagcacct ccttgcctc
 ctcccagctc tggattgca
 28441 gcttctctct ggctgcaaac tttctccaca atctaatgg
 aatgtcagtt tctcctgtt
 28501 cctgtccatc cgcaccact atcttcatgt tgttgagat
 gaagcgcgca agaccgtctg
 28561 aagatacctt caaccocgtg tatccatag acacggaaac
 cggctctcca actgtgcctt
 28621 ttcttactcc tccctttgta tcccccaatg ggtttcaaga
 gagtccccct ggggtactct

28681 ctttgcgcct atccgaacct ctagttagct ccaatggcat
 gcttgcgctc aaaatgggca
 28741 acggcctctc totggacgag gccggcaacc ttacctccca
 aatgtaacc actgtgagcc
 28801 cacctctcaa aaaaaccaag tcaaacataa acctggaaat
 atctgcaccc ctcacagtta
 28861 cctcagaagc cctaactgtg gctgccgccc cacctctaata
 ggtcgcgggc aacacactca
 28921 ccattgcaatc acaggccccg ctaaccgtgc acgactccaa
 acttagcatt gccacccaag
 28981 gaccctctac agtgtcagaa ggaaagctag ccctgcaaac
 atcaggcccc ctcaccacca
 29041 ccgatagcag tacccttact atcaactgct caccctctct
 aactactgcc actggttagct
 29101 tgggcattga cttgaaagag cccatttata cacaaaatgg
 aaaactagga ctaaagtacg
 29161 gggctccttt gcatgtaaca gacgacctaa acactttgac
 cgtagcaact ggtccaggtg
 29221 tgactattaa taatacttcc ttgcaacta aagttactgg
 agccttgggt tttgattcac
 29281 aaggcaatat gcaacttaat gtagcaggag gactaaggat
 tgattctcaa aacagacgcc
 29341 ttatacttga tgtagttat ccgtttgatg ctcaaaacca
 actaaatcta agactaggac
 29401 agggcctct ttttataaac tcagcccaca acttggatat
 taactacaac aaaggccttt
 29461 acttgtttac agcttcaaac aattccaaaa agcttgaggat
 taacctaagc actgccaagg
 29521 ggttgatggt tgacgctaca gccatagcca ttaatgcagg
 agatgggctt gaatttggtt
 29581 cacctaatgc accaaacaca aatcccctca aaacaaaaat
 tggccatggc ctagaatttg
 29641 attcaaaca ggctatggtt cctaaactag gaactggcct
 tagttttgac agcacaggtg
 29701 ccattacagt aggaaacaaa aataatgata agctaacttt
 gtggaccaca ccagctccat
 29761 ctctaactg tagactaat gcagagaaag atgctaaact
 cactttggtc ttaacaaaat
 29821 gtggcagtc aatacttgct acagtttcag ttttggctgt
 taaaggcagt ttggctccaa
 29881 tatctggaac agttcaaagt gctcatctta ttataagatt
 tgacgaaaat ggagtgtac
 29941 taacaatc cttcctggac ccagaatatt ggaactttag
 aatggagat cttactgaag
 30001 gcacagccta tacaacgct gttggattta tgcctaacct
 atcagcttat ccaaaatctc
 30061 acggtaaaac tgccaaaagt aacattgtca gtcaagttta
 cttaaacgga gacaaaacta
 30121 aacctgtaac actaaccatt aactaaacg gtacacagga
 aacaggagac acaactccaa

30181 gtgcatactc tatgtcattt tcatgggact ggtctggcca
 caactacatt aatgaaatat
 30241 ttgccacatc ctcttact tttcataca ttgcccaaga
 ataaagaatc gtttgtgtta
 30301 tgtttcaacg tgtttatfff tcaattgcag aaaatttcaa
 gtcatttttc attcagtagt
 30361 atagccccac caccacatag cttatacaga tcaccgtacc
 ttaatcaaac tcacagaacc
 30421 ctagtattca acctgccacc tccctcccaa cacacagagt
 acacagtcct ttctccccgg
 30481 ctggccttaa aaagcatcat atcatgggta acagacatat
 tcttaggtgt tatattccac
 30541 acggtttcct gtcgagccaa acgctcatca gtgatattaa
 taaactcccc gggcagctca
 30601 ctttaagttca tgtcgtctgc cagctgctga gccacaggct
 gctgtccaac ttgcgggtgc
 30661 ttaacgggcg gcgaaggaga agtccacgcc tacatggggg
 tagagtcata atcgtgcatc
 30721 aggatagggc ggtggtgctg cagcagcgcg cgaataaact
 gctgccgcg ccgctccgtc
 30781 ctgcaggaat acaacatggc agtggctctcc tcagcgatga
 ttgcaccgc ccgcagcata
 30841 aggcgccttg tctcctgggc acagcagcgc acctgatct
 cacttaaatc agcacagtaa
 30901 ctgcagcaca gcaccacaat attgttcaaa atcccacagt
 gcaaggcgct gtatccaaag
 30961 ctcatggcgg ggaccacaga acccagctgg ccatcatacc
 acaagcgcag gtagattaag
 31021 tggcgacccc tcataaacac gctggacata aacattacct
 cttttggcat gttgtaattc
 31081 accacctccc ggtaccatat aaacctctga ttaaacadgg
 cgccatccac caccatccta
 31141 aaccagctgg ccaaaacctg cccgccggct atacactgca
 gggaaccggg actggaacaa
 31201 tgacagtgga gagcccagga ctcgtaacca tggatcatca
 tgctcgtcat gatatcaatg
 31261 ttggcacaac acaggcacac gtgcatacac ttctcagga
 ttacaagctc ctcccgcgtt
 31321 agaaccatat cccagggaac aaccattcc tgaatcagcg
 taaatcccac actgcaggga
 31381 agacctcgca cgtaactcac gttgtgcatt gtcaaagtgt
 tacattcggg cagcagcgga
 31441 tgatcctcca gtatggtagc gcgggtttct gtctcaaaag
 gaggtagacg atccctactg
 31501 tacggagtgc gccgagacaa ccgagatcgt gttggtcgta
 gtgtcatgcc aatggaacg
 31561 ccggacgtag tcatatttcc tgaagcaaaa ccagggtgcg
 gcgtgacaaa cagatctgcg
 31621 tctccggtct cgcgcttag atcgtctctgt gtagtagttg
 tagtatatcc actctctcaa

31681 agcatccagg cgccccctgg cttcggggttc tatgtaaact
 ccttcatgcg cgcctgcctt
 31741 gataacatcc accaccgcag aataagccac acccagccaa
 cctacacatt cgttctgcga
 31801 gtcacacacg ggaggagcgg gaagagctgg aagaacatg
 tttttttttt tattccaaaa
 31861 gattatccaa aacctcaaaa tgaagatcta ttaagtgaac
 gcgctcccct cgggtggcgt
 31921 ggtcaaactc tacagccaaa gaacagataa tggcatttgt
 aagatgttgc acaatggctt
 31981 ccaaaaggca aacggccctc acgtccaagt ggacgtaaag
 gctaaaccct tcagggtgaa
 32041 tctcctctat aaacattcca gcaccttcaa ccatgcccaa
 ataattctca tctcggccacc
 32101 ttctcaatat atctctaagc aaatcccgaa tattaagtcc
 ggccattgta aaaatctgct
 32161 ccagagcgcc ctccacctc agcctcaage agcgaatcat
 gattgcaaaa attcaggttc
 32221 ctacagacc tgtataagat tcaaaagcgg aacattaaca
 aaaataccgc gatcccgtag
 32281 gtcccttcgc agggccagct gaacataatc gtgcaggctc
 gcacggacca gcgcggccac
 32341 ttccccgcca ggaaccttga caaaagaacc cacactgatt
 atgacacgca tactcggagc
 32401 tatgctaacc agcgtagccc cgatgtaage tttgttgcat
 gggcggcgat ataaaatgca
 32461 aggtgctgct caaaaaatca ggcaaagcct cgcgcaaaaa
 agaaagcaca tcgtagtcac
 32521 gctcatgcag ataaaggcag gtaagctccg gaaccaccac
 agaaaaagac accatttttc
 32581 tctcaaacat gtctgogggt ttctgcataa acacaaaata
 aaataacaaa aaaacattta
 32641 aacattagaa gcctgtctta caacaggaaa aacaaccctt
 ataagcataa gacggactac
 32701 ggccatgccg gcgtgaccgt aaaaaaactg gtcaccgtga
 ttaaaaagca ccaccgacag
 32761 ctctcgggtc atgtccggag tcataatgta agactcggta
 aacacatcag gttgattcat
 32821 cggtcagtgc taaaagcga ccgaaatagc ccgggggaat
 acatacccgc aggcgtagag
 32881 acaacattac agccccata ggaggtataa caaaattaat
 aggagagaaa aacacataaa
 32941 cacctgaaaa accctcctgc ctaggcaaaa tagcaccctc
 ccgctccaga acaacataca
 33001 gcgcttcaca gcggcagcct aacagtcagc cttaccagta
 aaaaagaaaa cctattaa
 33061 aacaccact cgacacggca ccagctcaat cagtcacagt
 gtaaaaaagg gccaaagtga
 33121 gagcgagtat atataggact aaaaaatgac gtaacgggta
 aagtcacaaa aaaacacca

33181 gaaaaccgca cgcgaaaccta cgcccagaaa cgaaagccaa
 aaaaccaca acttcctcaa
 33241 atcgtaactt cegttttccc acgttacgta actteccatt
 ttaagaaaac tacaattccc
 33301 aacacataca agttactccg ccctaaaacc tacgtcacc
 gcccggtcc cacgccccgc
 33361 gccacgtcac aaactccacc cctcattat catattggct
 tcaatccaaa ataaggtata
 33421 ttattgatga tnnnnnttaa t

SEQ ID NO: 17 (ND1.1 214)

1 taaggatccn nncctgtcct cgaccgatgc ccttgagagc cttcaacc
 gtcagctcct
 61 tccggtgggc gggggcatg actatcgctg ccgcacttat
 gactgtcttc tttatcatgc
 121 aactcgtagg acaggtgccg gcagcgtctt gggtcatttt
 cggcgaggac cgctttcgct
 181 ggagcgcgac gatgatcggc ctgtcgtctg cggatttcgg
 aatcttgcaac gccctcgtc
 241 aagccttcgt cactgggtccc gccaccaaac gtttcggcga
 gaagcaggcc attatcgccg
 301 gcatggcggc cgacgcgctg ggctacgtct tgctggcgtt
 cgcgacgcga ggctggatgg
 361 ccttccccat tatgattctt ctcgcttccg gggcatcgg
 gatgcccgcg tgcaggcca
 421 tgctgtccag gcaggtagat gacgaccatc agggacagct
 tcaaggatcg ctgcgggctc
 481 ttaccagcct aacttcgatc actggaccgc tgatcgtcac
 ggcgatttat gccgcctcgg
 541 cgagcacatg gaacggggtg gcatggattg taggcgccgc
 cctatacctt gtctgcctcc
 601 ccggttgcg tcgoggtgca tggagccggg ccacctcgac
 ctgaatggaa gccggcggca
 661 cctcgctaac ggattcacca ctccaagaat tggagccaat
 caattcttgc ggagaactgt
 721 gaatgcgcaa accaacctt ggcagaacat atccatcgcg
 tccgccatct ccagcagccg
 781 cacgcggcgc atctcgggca gcgttgggtc ctggccacgg
 gtgcgcatga tcgtgctcct
 841 gtcgttgagg acccggctag gctggcgggg ttgccttact
 ggtagcaga atgaatcacc
 901 gatagcgcgag cgaacgtgaa gcgactgctg ctgcaaaacg
 tctgcgacct gagcaacaac
 961 atgaatggtc ttcggtttcc gtgtttcgta aagtctggaa
 acgcggaagt cagcgcctg
 1021 caccattatg ttcggatct gcacgcagg atgctgctgg
 ctaccctgtg gaacacctac
 1081 atctgtatta acgaagcgtt ggcattgacc ctgagtgatt
 tttctctggc cccgcgcgat

1141 ccataccgcc agttgtttac cctcacaacg ttccagtaac
 cgggcatggt catcatcagt
 1201 aacccgatc gtgagcatcc tctctcgttt catcggtatc
 attacccccca tgaacagaaa
 1261 ttccccctta cacggaggca tcaagtgacc aaacaggaaa
 aaaccgccct taacatggcc
 1321 cgctttatca gaagccagac attaacgctt ctggagaaac
 tcaacgagct ggacgcggat
 1381 gaacaggcag acatctgtga atcgcttcac gaccacgctg
 atgagcttta cgcagctgc
 1441 ctcgcgctt tcggtgatga cggtgaaaac ctctgacaca
 tgcagctccc ggagacggtc
 1501 acagcttgtc tgtaagcggg tgccgggagc agacaagccc
 gtcagggcgc gtcagcgggt
 1561 gttggcgggt gtcggggcgc agccatgacc cagtcacgta
 gcgatagcgg agtgtatact
 1621 ggcttaacta tgcggcatca gagcagattg tactgagagt
 gcacccatag cgggtgtgaaa
 1681 taccgcacag atgcgtaagg agaaaatacc gcatcaggcg
 ctcttccgct tcctcgctca
 1741 ctgactcgct gcgctcggtc gttcggctgc ggcgagcggg
 atcagctcac tcaaaggcgg
 1801 taatacgggt atccacagaa tcaggggata acgcaggaaa
 gaacatgtga gcaaaaggcc
 1861 agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc
 gtttttccat aggetccgcc
 1921 ccctgacga gcatcaciaa aatcgacgct caagtcaag
 gtggcgaaac ccgacaggac
 1981 tataaagata ccaggcggtt cccctggaa gctccctcgt
 gcgctctcct gttccgaccc
 2041 tgccgcttac cggatacctg tccgccttcc tccttcggg
 aagcgtggcg ctttctcaat
 2101 gtcacgctg taggtatctc agttcgggtg aggtcgttcg
 ctccaagctg ggctgtgtgc
 2161 acgaaccccc cgttcagccc gaccgctgcg cttatccgg
 taactatcgt cttgagtcca
 2221 acccggtgag acacgactta tcgccactgg cagcagccac
 tggtaacagg attagcagag
 2281 cgaggtatgt aggcggtgct acagagttct tgaagtggg
 gcctaactac ggctacacta
 2341 gaaggacagt atttgggtatc tgcgctctgc tgaagccagt
 taccttcgga aaaagagttg
 2401 gtagctcttg atccggcaaa caaacaccg ctggtagcgg
 tggttttttt gtttgcaagc
 2461 agcagattac gcgcagaaaa aaaggatctc aagaagatcc
 tttgatcttt tctacggggg
 2521 ctgacgctca gtggaacgaa aactcacggt aagggatttt
 ggtcatgaga ttatcaaaaa
 2581 ggatcttcac ctagatcctt ttaaattaa aatgaagttt
 taaatcaatc taaagtatat

2641 atgagtaaac ttggtctgac agttaccaat gcttaatcag
 tgaggcacct atctcagcga
 2701 tctgtctatt tcgttcatcc atagttgcct gactccccgt
 cgtgtagata actacgatac
 2761 gggagggcct accatctggc cccagtgctg caatgatacc
 gcgagacca cgctcaccgg
 2821 ctccagattt atcagcaata aaccagccag ccggaagggc
 cgagcgcaga agtggtcctg
 2881 caactttatc cgctccatc cagtctatta attggtgccg
 ggaagctaga gtaagtagtt
 2941 cgccagttaa tagtttgcgc aacgttggtg ccattgctgc agccatga
 gattatcaaa
 aaggatcttc acctagatcc ttttcacgta gaaagccagt ccgcagaaaac
 ggtgctgacc
 ccgatgaat gtcagctact gggctatctg gacaagggaa aacgcaagcg
 caaagagaaa
 gcaggtagct tgcagtgggc ttacatggcg atagctagac tgggcggttt
 tatggacagc
 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag gttgggaagc
 cctgcaaagt
 aaactggatg gctttctcgc cgccaaggat ctgatggcgc aggggatcaa
 gctctgatca
 agagacagga tgaggatcgt ttcgcatgat tgaacaagat ggattgcacg
 caggttctcc
 ggccgcttgg gtggagagge tatteggcta tgactgggca caacagacia
 teggctgctc
 tgatgccgc gtgttccggc tgcagcgcga ggggcgcccg gttctttttg
 tcaagaccga
 cctgtccggg gccctgaatg aactgcaaga cgaggcagcg cggctatcgt
 ggetggccac
 gacgggcggt ccttgcgcag ctgtgctcga cgttgctact gaagcgggaa
 gggactggct
 gctattgggc gaagtgccgg ggcaggatct cctgtcatct caccttgctc
 ctgccgagaa
 agtatccatc atggctgatg caatgcggcg gctgcatacg cttgatccgg
 ctacctgcc
 attcgaccac caagcgaaac atcgcacgca gcgagcacgt actcggatgg
 aagccggctc
 tgtcgatcag gatgatctgg acgaagagca tcaggggctc gcgccagccg
 aactgttcgc
 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc gtgacctatg
 gcgatgcctg
 cttgccgaat atcatggtgg aaaatggcgg cttttctgga ttcatcgact
 gtggccggct
 ggggtgtggcg gaccgctatc aggacatagc gttggctacc cgtgatattg
 ctgaagagct
 tggcgcgcaa tgggctgacc gcttccctgt gctttacggg atcgcgcctc
 ccgattcgca
 gcgcacgcgc ttctatcgcc ttcttgacga gttcttctga attttgtaa
 aatttttggt

aatcagctc atttttaac caataggccg aatcggcaa catcccttat
 aatcaaaag
 aatagaccgc gatagggttg agtggtgttc cagtttgaa caagagtcca
 ctattaaaga
 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc
 ccactacgtg
 aaccatcacc caaatcaagt tttttgcggt cgaggtgccg taaagctcta
 aatcgggaacc
 ctaaaggag cccccgattt agagcttgac ggggaaagcc ggcgaaactg
 gcgagaaagg
 aagggaagaa agcgaaagga gggggcgcta gggcgctggc aagtgtagcg
 gtcacgctgc
 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag ggcgctcca
 ttcgccattc
 aggatcgaat taattcttaa ttaacatcacc aataatatac cttattttgg
 attgaagcca atatgataat gagggggttg agtttgtgac gtggcgoggg
 gcgtgggaac ggggcgggtg acgtagtagt gtggcggaag tgtgatgttg
 caagtgtggc ggaacacatg taagcgacgg atgtggcaaa agtgacgttt
 ttggtgtgcg ccggtgtaca caggaagtga caattttcgc gcggttttag
 gcggatgttg tagtaaattt gggcgtaacc gagtaagatt tggccatttt
 cgcgggaaaa ctgaataaga ggaagtgaaa tctgaataat tttgtgttac
 tcatagcgcg
 taatactGCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC
 CAGGGTFTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
 TAGGGCGAATTGGGTACTGGCCAacagagcttggcccattgcatacgttgtatccatatacata
 atatgtacatttatattgggtcatgtccaacattaccgccatgttgacattgattattgact
 agttattaatagtaatcaattaacggggtcattagttcatagcccatatattggagttccgcgt
 tacataacttacggtaaatggccgcctggctgaccgcccacgacccccgcccattgacgt
 caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtatttacggtaaaactgccacttggcagtagatcaatgtagtcatatgccaagtacgcc
 cctattgacgtcaatgacggtaaatggccgcctggcattatgccagtagacatgacctat
 gggactttcctacttggcagtagatctacgtattagtagtcatcgctattaccatgggtgatgcgg
 ttttggcagtagatcaatgggcgtggatagcggtttgactcacggggatttccaagtctcca
 cccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtc
 gtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgtcagatcgctggagacgccatccacgctgttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttggtggtgag
 gccctgggcaggttgggtatcaaggttacaagacaggtttaaggagaccaatagaaactgggc
 atgtggagacagagaagactccttgggtttctgataggcactgactctctctgcctattgggtc
 tattttcccacccttaggctgctggtctgagcctagGAGATCTCTCGAGGTCGACGGTATcg
 atgcca cc**ATGGAGAA AATCGTCCTG TTGCTCGCTA TTGTGTCTCTAGTGAAGAGC**
GATCAAATTT GATCGGCTA CCATGCCAAT AACTCAACAGAGCAGGTCGA
TACTATCATG GAGAAAAACG TAACAGTTAC TCATGCCCAA GACATCTTGG
AAAAGACCCA CAACGGCAAA CTTTGCGACC TGGATGGAGT GAAGCCCCTG
ATCCTCCGGG ACTGTTTAGT CGCTGGTTGG CTGCTCGGGA ACCCTATGTG
TGATGAGTTT ATCAACGTGC CTGAATGGTC TTACATTGTG GAGAAGGCTA
ACCCTACCAA TGACCTCTGC TATCCTGGGT CATTTAACGA TTACGAGGAA
CTGAAACACC TGTTGTCTAG AATTAACCAC TTTGAAAAGA TACAGATTAT
ACCCAAGTCT AGTTGGAGTG ATCACGAAGC CTCCTCAGG GTTAGCTCAG
CGTGTCCTA TCTGGGCTCT CCATCCTTCT TTAGAAATGT GGTCTGGTTA
ATCAAAAAGA ACAGTACCTA CCCAACCATC AAAAAGTCTT ATAACAATAC

CAATCAGGAG GACCTGCTCG TGTTGTGGGG TATCCATCAC CCGAACGACG
CCGCTGAACA GACTAGGCTG TATCAGAACC CCACTACATA CATCAGTATT
GGCACGAGTA CTCTGAACCA GCGATTAGTG CCAAAGATTG CAACACGGAG
CAAAGTAAAT GGGCAATCTG GCAGGATGGA GTTTTTCTGG ACAATCTTAA
AACCCAACGA TCGGATAAAT TTCGAGTCCA ATGGCAATTT CATCGCCCCT
GAATACGCCT ATAAGATCGT GAAAAAGGGG GACTCTGCAA TTATGAAGTC
CGAATTAGAG TATGGCAATT GCAACACGAA GTGCCAGACA CCAATGGGAG
CCATTAATAG CTCAATGCCC TTCCATAATA TTCATCCATT GACCATTGGG
GAGTGCCCAA AGTACGTGAA GTCCAACCGC CTGGTCCTCG CAACCGGTCT
AAGAAATAGC CCGCAGAGAG AATCGCGGAG GAAGAAACGT GGCCTGTTTG
GCGCGATTGC CGGATTCATC GAGGGAGGCT GGCAGGGTAT GGTTCGATGGT
TGGTACGGAT ACCACCATAG CAACGAACAG GGGTCCGGCT ATGCAGCAGA
TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAAC AAGGTTAATA
GTATTATAGA TAAAATGAAC ACGCAATTCG AGGCCGTTGG GAGGGAGTTT
AACAATCTGG AACGCCGGAT CGAAAATCTG AATAAGAAAA TGGAAGACGG
CTTCCTTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA CTCATGGAGA
ACGAGAGGAC CCTGGATTC CACGATAGCA ACGTGAAGAA CCTTTACGAC
AAGGTGAGAC TTCAGCTCCG AGACAACGCC AAGGAGCTGGGGAATGGATG
CTTCGAGTTT TACCACAAAT GTGACAATGA GTGCATGGAAAGTATACGCA
ACGGGACCTA CAATTACCCT CAGTATAGCG AAGAGGCTCGGCTCAAACGC
GAAGAGATAA GCGGGGTGAA ATTGGAATCA ATCGGAACATATCAAATCCT
GTCCATCTAT TCCACCGTCG CCTCTTCGCT GGCCCTCGCT
ATCATGATGG CTGGTCTGTC CCTATGGATG TGTTCCAATG
GAAGCCTTCAGTGCCGTATT TGTATATGAg c
GGCCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCC
TT
CTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
ACTCCCAC
TGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
TGGGGGGT
GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
GGTGGGCT
CTATGGCTTCTGAGGCGGAAAGAACCAAAGCTtaacatcatc aataatatac
cttatttttgg attgaagcca atatgataat gagggggtgg.
agtttgtgac gtggcgcggg gcgtgggaac ggggcggggtg acgtagtagt
gtggcggaag
tgtgatgttg caagtgtggc ggaacacatg taagcgacgg atgtggcaaa
agtgacgttt
ttggtgtgcg cgggtgtaca caggaagtga caattttcgc gcggttttag
gcggatgttg
tagtaaattt gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa
ctgaataaga
ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg taatactgta
atagtaatca
attacgggggt cattagttca tagccatat atggagttcc gcgttacata
acttacggta
aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat
aatgacgtat
gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga
gtatttacgg

ES 2 573 456 T3

taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc
 ccctattgac
 gtcaatgacg gtaaattggcc cgcctggcat tatgcccagt acatgacctt
 atgggacttt
 cctacttggc agtacatcta cgtattagtc atcgctatta ccatgggtgat
 gcggttttgg
 cagtacatca atgggcgtgg atagcgggtt gactcacggg gatttccaag
 tctccacccc
 attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc
 aaaatgtcgt
 aacaactccg cccattgac gcaaatgggc ggtaggcgtg tacgggtggga
 ggtctatata
 agcagagctg gtttagtgaa ccgtcagatc
 cgctagAGATCTGGgaaacgatatgggctgaatac**ggatcc**gatttcagcccatatcgtttc
 TCT AGA AAT AAA ATA TCT TTA TTT TCA TTA CAT CTG TGT GTT GGT
 TTT TTG TGT GGC
 GG CCGCtcgagcct AAGCTTctag ataagatata cgatccaccg gatctagata
 actgatcata
 atcagccata ccacatttgt agaggtttta ctgtcttaa aaaacctccc
 acacctccc
 ctgaacctga aacataaaat gaatgcaatt gttgttgta acttgtttat
 tgcagcttat
 aatggttaca aataaagcaa tagcatcaca aatttcacaa ataaagcatt
 tttttcactg
 cattctagtt gtggtttgtc caaactcacc aatgtatctt aacgcggatc
 tgggcgtggt
 taagggtggg aaagaatata taagggtggg gtcttatgta gttttgtatc
 tgttttgcag
 cagccgccgc cgccatgagc accaactcgt ttgatggaag cattgtg
 AGCTTgtcgactcg aagatctggg cgtgggtaag
 ggtgggaaag aatatataag gtgggggtct tatgtagttt tgtatctggt
 ttgcagcagc
 cgccgccgcc atgagcacca actcgtttga tgggaagcatt gtgagctcat
 atttgacaac
 gcgcatgccc ccatgggccg ggggtgcgtca gaatgtgatg ggctccagca
 ttgatgggtc
 cccgctcctg cccgcaaact ctactacctt gacctacgag accgtgtctg
 gaacgcggt
 ggagactgca gctccgccg ccgcttcagc cgctgcagcc accgcccgcg
 ggattgtgac
 tgactttgct ttctgagcc cgcttgcaag cagtgcagct tcccgttcat
 cogcccgcga
 tgacaagttg acggctcttt tggcacaatt ggattctttg acccgggaac
 ttaatgtcgt
 ttctcagcag ctgttggatc tgcgccagca ggtttctgcc ctgaaggctt
 cctccccctc
 caatgoggtt taaaacataa ataaaaaac agactctggt tggatttggg
 tcaagcaagt
 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag gcccgggacc
 agcgggtctc

gtcggtgagg gtcctgtgta tttttccag gacgtggtaa aggtgactct
 ggatgttcag
 atacatgggc ataagcccgt ctctgggggtg gaggtagcac cactgcagag
 cttcatgctg
 cggggtggtg ttgtagatga tocagtcgta gcaggagcgc tgggctggtg
 gcctaaaaat
 gtctttcagt agcaagctga ttgccagggg caggcccttg gtgtaagtgt
 ttacaaagcg
 gttaagctgg gatgggtgca tacgtgggga tatgagatgc atcttgact
 gtatttttag
 gttggctatg ttcccagcca tatccctccg gggattcatg ttgtgcagaa
 ccaccagcac
 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta gaaggaaatg
 cgtggaagaa
 cttggagacg ccctgtgac ctccaagatt ttccatgcat tcgtccataa
 tgatggcaat
 gggcccacgg gcggcggcct gggcgaagat atttctggga tcaactaacgt
 catagttgtg
 ttccaggatg agatcgatc aggccatttt tacaagcgc gggcggaggg
 tgccagactg
 cggtataatg gttccatccg gcccaggggc gtagttacc tccagatgtt
 gcatttccca
 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg gcgatgaaga
 aaacggttc
 cggggtaggg gagatcagct ggaagaaag caggttcctg agcagctgcg
 acttaccgca
 gccggtgggc ccgtaaatca cacctattac cgggtgcaac tggtagttaa
 gagagct
 g cagctgccgt catccctgag caggggggccc
 acttcgttaa gcatgtccct gactcgcag ttttccctga ccaaattccg
 cagaaggcgc
 tcgccgccca gcgatagcag ttcttgcaag gaagcaaagt ttttcaacgg
 tttgagaccg
 tccgccgtag gcatgctttt gagcgtttga ccaagcagtt ccaggcggtc
 ccacagctcg
 gtcacctgct ctacggcacc tcgatccagc atatctcctc gtttcgcggg
 ttggggcggc
 tttcgcgtga cggcagtagt cgggtgctcgt ccagacgggc cagggtcag
 tctttccacg
 ggcgcagggt cctcgtcagc gtagtctggg tcacggtgaa ggggtgcgct
 ccgggctgcg
 cgctggccag ggtgcgcttg aggctggctc tgctggtgct gaagcgcctg
 cggctctcgc
 cctgcgcgct gcccaggtag catttgacca tggtgtcata gtccagcccc
 tccgcggcgt
 ggcccttggc gcgcagcttg cccttgagg aggccgccca cgaggggag
 tgcagacttt
 tgagggcgta gagcttgggc gcgagaaata ccgattccgg ggagtaggca
 tccgcgcgcg
 aggccccgca gacggctctc cattccacga gccaggtgag ctctggccgt
 tcggggtcaa

aaaccagggt tccccatgc tttttgatgc gtttettacc tctggtttcc
 atgagccggt
 gtccacgctc ggtgacgaaa aggctgtccg tgtccccgta tacagacttg
 agaggcctgt
 cctcgagcgg tgttccgcgg tctctctcgt atagaaactc ggaccactct
 gagacaaagg
 ctgcgctcca ggccagcacg aaggaggcta agtgggaggg gtagcggctc
 ttgtccacta
 gggggtccac tcgctccagg gtgtgaagac acatgtcggc ctcttcggca
 tcaaggaagg
 tgattggttt gtaggtgtag gccacgtgac cgggtgttcc tgaagggggg
 ctataaaagg
 ggggtggggc gcgttcgtcc tcactctctt ccgcatcgtc gtctgcgagg
 gccagctgtt
 ggggtgagta ctccctctga aaagcgggca tgacttctgc gctaagattg
 tcagtttcca
 6301 aaaacgagga ggatttgata ttcacctggc ccgcggtgat
 gcctttgagg gtggccgcat
 6361 ccatctggtc agaaaagaca atctttttgt tgtcaagctt
 ggtggcaaac gaccctaga
 6421 gggcgttgga cagcaacttg gcgatggagc gcagggtttg
 gtttttgtcg cgatcggcgc
 6481 gctccttggc cgcgatgttt agctgcacgt attcgcgcgc
 aacgcaccgc cattcgggaa
 6541 agacggtggt gcgctcgtcg ggcaccaggc gcacgcgcca
 accgcggttg tgcaggggta
 6601 caaggtcaac gctggtgget acctctccgc gtaggcgctc
 gttggtccag cagaggcggc
 6661 cgcccttgcg cgagcagaat ggcggtaggg ggtctagctg
 cgtctcgtcc ggggggtctg
 6721 cgtccacggt aaagaccccg ggcagcaggc gcgcgtcga
 gtagtctatc ttgcatcctt
 6781 gcaagtctag cgcctgctgc catgcgcggg cggcaagcgc
 gcgctcgtat gggttgagt
 6841 ggggaccca tggcatgggg tgggtgagcg cggaggcgta
 catgccgcaa atgtcgtaaa
 6901 cgtagagggg ctctctgagt attccaagat atgtagggta
 gcatcttcca ccgcggatgc
 6961 tggcgcgcac gtaatcgtat agttcgtcgc agggagcgcg
 gaggtcggga ccgaggttgc
 7021 tacgggcggg ctgctctgct cgggaagacta tctgcctgaa
 gatggcatgt gagttggatg
 7081 atatggttgg acgctggaag acggttgaagc tggcgtctgt
 gagacctacc gcgtcacgca
 7141 cgaaggaggc gtaggagtcg cgcagcttgt tgaccagctc
 ggcggtgacc tgcacgtcta
 7201 gggcgcagta gtccagggtt tccttgatga tgtcatactt
 atcctgtccc tttttttcc
 7261 acagctcgcg gttgaggaca aactcttcgc ggtctttcca
 gtactcttgg atcggaaacc

7321 cgtcggcctc ogaacggtaa ggcctagca tgtagaactg
 gttgacggcc tggtaggcgc
 7381 agcatccctt ttctaagggt agcgcgatg cctgcgcggc
 cttccggagc gaggtgtggg
 7441 tgagcgcaaa ggtgtccctg accatgactt tgaggtagctg
 gtatttgaag tcagtgtcgt
 7501 cgcacccgcc ctgctcccag agcaaaaagt ccgtgcgctt
 ttggaacgc ggatttggca
 7561 gggcgaagg gacatcgttg aagagtatct ttcccgcgcg
 aggcataaag ttgcgtgtga
 7621 tgcggaaggg tcccggcacc tcggaacggg tgtaattac
 ctgggcccgc agcacgatct
 7681 cgtcaaagcc gttgatggtg tggcccacaa tgtaaagttc
 caagaagcgc gggatgccct
 7741 tgatggaagg caatttttta agttcctcgt aggtgagctc
 ttcaggggag ctgagcccgt
 7801 gctctgaaag ggcccagtct gcaagatgag ggttgaagc
 gacgaatgag ctccacaggc
 7861 cacgggccat tagcatttgc aggtggcgc gaaaggtcct
 aaactggcga cctatggcca
 7921 tttttctgg ggtgatgcag tagaaggtaa gcggtcttg
 ttcccagcgg tcccatccaa
 7981 ggttcgcggc taggtctcgc gcggcagtc ctagaggctc
 atctccgccc aacttcatga
 8041 ccagcatgaa gggcacgagc tgcttcccaa aggccccat
 ccaagtatag gtctctacat
 8101 cgtaggtgac aaagagagc tcggtgcgag gatgcgagcc
 gatcgggaag aactggatct
 8161 cccgccacca attggaggag tggctattga tgtggtgaaa
 gtagaagtcc ctgcgacggg
 8221 ccgaacaact gtgctggctt ttgtaaaaac gtgcgcagta
 ctggcagcgg tgcaocggct
 8281 gtacatcctg cacgaggttg acctgacgac cgcgcacaag
 gaagcagagt gggaatttga
 8341 gccctcgc tggcgggtt ggctgggtgt cttctacttc
 ggctgcttgt ccttgaccgt
 8401 ctggctgctc gaggggagtt acggtggatc ggaccaccac
 gccgcgcgag ccaaagtc
 8461 agatgtccgc gcgcggcggc cggagcttga tgacaacatc
 gcgcagatgg gagctgtcca
 8521 tggctcggag ctcccgcggc gtcaggtcag gcgggagctc
 ctgcaggttt acctcgcata
 8581 gacgggtcag ggcgcgggct agatccaggt gataccta
 ttccaggggc tggttggtg
 8641 cggcgtcgt ggcttgcaag aggccgcac cccgcggcgc
 gactacggta ccgcgcggc
 8701 ggcggtgggc cgcgggggtg tccttggtat atgcatctaa
 aagcgtgac gcgggcgagc
 8761 ccccggaggt agggggggct ccggaccgc cgggagaggg
 ggcaggggca cgtcggcgc

8821 gcgcgcgggc aggagctggt gctgcgcgcg tagggtgctg
 gcgaacgcga cgacgcggcg
 8881 gttgatctcc tgaatctggc gcctctgcgt gaagacgcgc
 ggcccgggtga gcttgagcct
 8941 gaaagagagt tcgacagaat caatttcggg gtcggtgacg
 gcggcctggc gcaaaatctc
 9001 ctgcacgtct cctgagttgt cttgataggc gatctcggcc
 atgaactgct cgatctcttc
 9061 ctccctggaga tctccgcgtc cggctcgcctc cacgggtggcg
 gcgaggtcgt tggaaatgcg
 9121 ggccatgagc tgcgagaagg cgttgaggcc tcctcgttc
 cagacgcggc tgtagaccac
 9181 gcccccttcg gcatcgcggg cgcgcatgac cacctgcgcg
 agattgagct ccacgtgccg
 9241 ggcgaagacg gcgtagtttc gcaggcgcctg aaagaggtag
 ttgaggggtgg tggcgggtgtg
 9301 ttctgccacg aagaagtaca taaccacgcg tcgcaacgtg
 gattcgttga tatcccccaa
 9361 ggctcaagg cgctccatgg cctcgtagaa gtccacggcg
 aagttgaaaa actgggagtt
 9421 gcgcgcgcgc acgggttaact cctcctccag aagacggatg
 agctcggcga cagtgtcgcg
 9481 cacctcgcgc tcaaaggcta caggggcctc ttcttcttct
 tcaatctcct ctccataag
 9541 ggctccctt tcttcttctt ctggcggcgg tgggggaggg
 gggacacggc ggcgacgcg
 9601 gcgcaccggg aggcggtcga caaagcgcctc gatcatctcc
 ccgcggcgac ggcgcatggt
 9661 ctcggtgacg gcgcggccgt tctcgcgggg gcgcagttgg
 aagacgccgc ccgtcatgtc
 9721 ccggttatgg gttggcgggg ggctgccatg cggcagggat
 acggcgctaa cgatgcatct
 9781 caacaattgt tgtgtaggta ctccgcgcgc gagggacctg
 agcgagtccg catcgaccgg
 9841 atcggaaaac ctctcgagaa aggcgtctaa ccagtcacag
 tcgcaaggta ggctgagcac
 9901 cgtggcgggc ggcagcgggc ggcggtcggg gttgtttctg
 gcggaggtgc tgctgatgat
 9961 gtaattaaag taggcgggtct tgagacggcg gatggtogac
 agaagcacca tgccttggg
 10021 tccggcctgc tgaatgcgca ggcggtcggc catgccccag
 gcttcgtttt gacatcggcg
 10081 caggtctttg tagtagtctt gcatgagcct ttctaccggc
 acttcttctt ctcttctc
 10141 ttgtcctgca tctcttgcat ctatcgtctg gcggcgggcg
 gagtttggcc gtaggtggcg
 10201 ccctcttct cccatgcgtg tgacccccgaa gccctcctc
 ggctgaagca gggctaggtc
 10261 ggcgacaacg cgctcggcta atatggcctg ctgcacctgc
 gtgagggtag actggaagtc

10321 atccatgtcc acaaagcggg ggtatgcgcc cgtggtgatg
 gtgtaagtgc agttggccat
 10381 aacggaccag ttaacggtct ggtgacccgg ctgcgagagc
 tcggtgtacc tgagacgcga
 10441 gtaagccctc gagtcaaata cgtagtcggt gcaagtccgc
 accaggtact ggtatcccac
 10501 caaaaagtgc ggcggcggct ggcggtagag gggccagcgt
 aggggtggccg gggctccggg
 10561 ggcgagatct tccaacataa ggcgatgata tccgtagatg
 tacctggaca tccaggtgat
 10621 gccggcggcg gtggtggagg cgcgcggaaa gtcgcggacg
 cggttccaga tgttgccgag
 10681 cggcaaaaag tgctccatgg tcgggacgct ctggccggtc
 aggcgcgcgc aatcgttgac
 10741 gctctaccgt gcaaaaggag agcctgtaag cgggcactct
 tccgtggtct ggtggataaa
 10801 ttcgcaaggg tatcatggcg gacgaccggg gttcgagccc
 cgtatccggc cgtccgccc
 10861 gatccatgcg gttaccgccc gcgtgtcgaa cccaggtgtg
 cgacgtcaga caacggggga
 10921 gtgctccttt tggcttctt ccaggcgcgg cggctgctgc
 gctagctttt ttggccactg
 10981 gccgcgcgca gcgtaagcgg ttaggctgga aagcgaagc
 attaagtggc tcgctccctg
 11041 tagccggagg gttatthtcc aagggttgag tcgcgggacc
 cccggttcga gtctcggacc
 11101 ggccggactg cggcgaacgg gggtttgcct ccccgatcg
 caagaccocg cttgcaaatt
 11161 cctccggaaa cagggacgag cccctttttt gcttttcca
 gatgcatccg gtgctgcggc
 11221 agatgcgccc ccctcctcag cagcggcaag agcaagagca
 gcggcagaca tgcagggcac
 11281 cctcccctcc tcctaccgcg tcaggagggg cgacatccgc
 ggttgacgcg gcagcagatg
 11341 gtgattacga acccccgcgg cgcggggccc ggcactacct
 ggacttgag gagggcgagg
 11401 gcctggcgcg gctaggagcg ccctctcctg agcggtaacc
 aagggtgcag ctgaagcgtg
 11461 atacgcgtga ggcgtacgtg ccgcggcaga acctgtttcg
 cgaccgcgag ggagaggagc
 11521 ccgaggagat gcgggatcga aagttccacg cagggcgcga
 gctgcggcat ggctgaatc
 11581 gcgagcggtt gctgcgcgag gaggactttg agcccgacgc
 gcgaaccggg attagtcccg
 11641 cgcgcgcaca cgtggcggcc gccgacctgg taaccgcata
 cgagcagacg gtgaaccagg
 11701 agattaactt tcaaaaaagc ttaacaacc acgtgcgtac
 gcttgtggcg cgcgaggagg
 11761 tggctatagg actgatgcat ctgtgggact ttgtaagcgc
 gctggagcaa aaccacaata

11821 gcaagccgct catggcgcag ctgttcctta tagtgcagca
 cagcagggac aacgagggcat
 11881 tcagggatgc gctgctaaac atagtagagc ccgagggccg
 ctggctgctc gatttgataa
 11941 acatcctgca gagcatagtg gtgcaggagc gcagcttgag
 cctggctgac aaggtggccg
 12001 ccatcaacta ttccatgctt agcctgggca agttttacgc
 ccgcaagata taccataccc
 12061 cttacgttcc catagacaag gaggtaaaga tcgaggggtt
 ctacatgcgc atggcgctga
 12121 aggtgcttac cttgagcgcac gacctgggcg tttatcgcaa
 cgagcgcac cacaaggccg
 12181 tgagcgtgag ccggcggcgc gagctcagcg accgcgagct
 gatgcacagc ctgcaaaggg
 12241 ccctggctgg cacgggcagc ggcgatagag aggccgagtc
 ctactttgac gcgggcgctg
 12301 acctgcgctg ggccccaagc cgacgcgccc tggaggcagc
 tggggccgga cctgggctgg
 12361 cgggtggcacc cgcgcgcgct ggcaacgtcg gcggcgtgga
 ggaatatgac gaggacgatg
 12421 agtacgagcc agaggacggc gagtactaag cggatgatgtt
 tctgatcaga tgatgcaaga
 12481 cgcaacggac ccggcgggtgc gggcggcgcct gcagagccag
 ccgtccggcc ttaactccac
 12541 ggacgactgg cgccagggtca tggaccgcat catgtcgcctg
 actgcgcgca atcctgacgc
 12601 gttccggcag cagccgcagc ccaaccggct ctccgcaatt
 ctggaagcgg tggccccggc
 12661 gcgcgcaaac ccacgcacg agaaggtgct ggcgatcgta
 aacgcgctgg ccgaaaacag
 12721 ggccatccgg cccgacgagg ccggcctggt ctacgacgcg
 ctgcttcagc gcgtggctcg
 12781 ttacaacagc ggcaacgtgc agaccaacct ggaccggctg
 gtgggggatg tgccgcgagc
 12841 cgtggcgcag cgtgagcgcg cgcagcagca gggcaacctg
 ggctccatgg ttgcactaaa
 12901 cgcttcctg agtacacagc ccgccaacgt gccgcgggga
 caggaggact acaccaactt
 12961 tgtgagcgc cactgaggctaa tggatgactga gacaccgcaa
 agtgaggtgt accagtctgg
 13021 gccagactat tttttccaga ccagtagaca aggccctgcag
 accgtaaacc tgagccagc
 13081 tttcaaaaac ttgcaggggc tgtggggggt gcgggctccc
 acaggcgacc gcgcgaccgt
 13141 gtctagcttg ctgacgccca actcgcgcct gttgctgctg
 ctaatagcgc cttcacgga
 13201 cagtggcagc gtgtcccggg acacatacct aggtcacttg
 ctgacactgt accgcgagc
 13261 cataggtcag gcgcacgtgg acgagcatac tttccaggag
 attacaagtg tcagccgcgc

13321 gctggggcag gaggacacgg gcagcctgga ggcaacccta
 aactacctgc tgaccaaccg
 13381 gcggcagaag atcccctcgt tgcacagttt aaacagcgag
 gaggagcgca ttttgcgcta
 13441 cgtgcagcag agcgtgagcc ttaacctgat gcgcgacggg
 gtaacgcccc gcgtggcgct
 13501 ggacatgacc gcgcgcaaca tggaaccggg catgtatgcc
 tcaaaccggc cgtttatcaa
 13561 ccgcctaata gactacttgc atcgcgcggc cgccgtgaac
 cccgagtatt tcaccaatgc
 13621 catcttgaac ccgcaactggc taccgcccc tggtttctac
 accgggggat tcgaggtgcc
 13681 cgagggtaac gatggattcc tctgggacga catagacgac
 agcgtgtttt ccccgcaacc
 13741 gcagaccctg ctagagtgc aacagcgcgga gcaggcagag
 gcggcgctgc gaaaggaaag
 13801 cttccgcag ccaagcagct tgtccgatct aggcgctgcg
 gccccgcggt cagatgctag
 13861 tagccattt ccaagcttga tagggtctct taccagcact
 cgcaccacc gcccgcgct
 13921 gctgggcgag gaggagtacc taacaactc gctgctgcag
 ccgcagcgcg aaaaaaacct
 13981 gcctccggca tttcccaaca acgggataga gagcctagtg
 gacaagatga gtagatggaa
 14041 gacgtacgcg caggagcaca gggacgtgcc aggcccgcgc
 ccgcccacc gtcgtcaaag
 14101 gcacgaccgt cagcggggtc tgggtgaggga ggacgatgac
 tcggcagacg acagcagcgt
 14161 cctggatttg ggagggagtg gcaaccggt tgcgcacctt
 cgccccaggc tggggagaat
 14221 gttttaaaa aaaaaagca tgatgcaaaa taaaaactc
 accaaggcca tggcaccgag
 14281 cgttggttt cttgtattcc ccttagtatg cggcgcgcgg
 cgatgtatga ggaaggtcct
 14341 cctccctcct acgagagtgt ggtgagcgcg gcgccagtgg
 cggcggcgct gggttctccc
 14401 ttcgatgctc ccctggacc gccgtttggt cctccgcgg
 acctgcggcc taccggggg
 14461 agaaacagca tccgttactc tgagttggca cccctattcg
 acaccaccg tgtgtacctg
 14521 gtggacaaca agtcaacgga tgtggcatcc ctgaactacc
 agaacgacca cagcaactt
 14581 ctgaccacgg tcattcaaaa caatgactac agcccggggg
 aggcaagcac acagaccatc
 14641 aatcttgacg accggtcgca ctggggcggc gacctgaaaa
 ccatcctgca taccaacatg
 14701 ccaaatgtga acgagttcat gtttaccat aagttaagg
 cgcggtgat ggtgtcgcg
 14761 ttgcctacta aggacaatca ggtggagctg aaatacagat
 ggggtgagtt cacgctgcc

14821 gagggcaact actccgagac catgaccata gaccttatga
 acaacgcgat cgtggagcac
 14881 tacttgaaag tgggcagaca gaacgggggt ctggaaagcg
 acatcgggggt aaagtttgac
 14941 acccgcaact tcagactggg gtttgacccc gtcactggtc
 ttgtcatgcc tggggtatat
 15001 acaaacgaag ccttccatcc agacatcatt ttgctgccag
 gatgcgggggt ggacttcacc
 15061 cacagccgcc tgagcaactt gttgggcac c gcaagcggc
 aacccttoca ggagggcttt
 15121 aggatcacct acgatgatct ggagggtggt aacattoccg
 cactgttggg tgtggacgcc
 15181 taccaggcga gcttgaaaga tgacaccgaa cagggcgggg
 gtggcgcagg cggcagcaac
 15241 agcagtggca gcggcgcgga agagaactcc aacgcggcag
 ccgcggcaat gcagccgggtg
 15301 gaggacatga acgatcatgc cattcgcggc gacacctttg
 ccacacgggc tgaggagaag
 15361 cgcgctgagg ccgaagcagc ggccgaagct gcgcgccccg
 ctgcgcaacc cgaggtcgag
 15421 aagcctcaga agaaaccggt gatcaaacc ctgacagagg
 acagcaagaa acgcagttac
 15481 aacctaataa gcaatgacag caccttcacc cagtaccgca
 gctggtacct tgcatacaac
 15541 tacggcgacc ctgagaccgg aatccgctca tggaccctgc
 tttgactec tgacgtaacc
 15601 tgcggctcgg agcaggctca ctggtcgttg ccagacatga
 tgcaagaccc cgtgaccttc
 15661 cgctccacgc gccagatcag caactttccg gtggtgggcg
 ccgagctggt gcccgctcac
 15721 tccaagagct tctacaacga ccaggccgct tactcccaac
 tcatccgcca gtttacctct
 15781 ctgaccacag tgttcaatcg ctttcccgag aaccagattt
 tggegcgccc gccagcccc
 15841 accatcacca ccgtcagtga aaacgttcct gctctcacag
 atcacgggac gctaccgctg
 15901 cgcaacagca tcggaggagt ccagcgagtg accattactg
 acgccagacg ccgcacctgc
 15961 ccctacgttt acaaggccct gggcatagtc tcgcccgcgcg
 tcctatcgag ccgcactttt
 16021 tgagcaagca tgtccatcct tatatcgccc agcaataaca
 caggctgggg cctgcgcttc
 16081 ccaagcaaga tgtttggcgg ggccaagaag cgctccgacc
 aacacccagt gcgcgtgcgc
 16141 gggcactacc gcgcgcctg gggcgcgcac aaacgcggcc
 gcactgggcg caccaccgct
 16201 gatgacgcca tcgacgcggt ggtggaggag gcgcgcaact
 acacgcccac gccgccacca
 16261 gtgtccacag tggacgcggc cattcagacc gtggtgcgcg
 gagcccggcg ctatgctaaa

16321 atgaagagac ggcggaggcg cgtagcacgt cgccaccgcc
 gccgaccgg cactgccgcc
 16381 caacgcgcg cggcggcct gcttaaccgc gcacgtcgca
 ccggccgacg ggcggccatg
 16441 cgggccgctc gaaggctggc cgoggggtatt gtcactgtgc
 cccccaggtc caggcgacga
 16501 gcggccgccc cagcagccgc ggccattagt gctatgactc
 agggtcgcag gggcaacgtg
 16561 tattgggtgc gcgactcggg tagcggcctg cgcggtgccc
 tgccgacccc cccccgcgc
 16621 aactagattg caagaaaaaa ctacttagac tcgtactgtt
 gtatgtatcc agcggcggcg
 16681 gcgcgcaacg aagctatgtc caagcgcaaa atcaaagaag
 agatgctcca ggtcatcgcg
 16741 ccggagatct atggcccccc gaagaaggaa gagcaggatt
 acaagccccg aaagctaaag
 16801 cgggtcaaaa agaaaaagaa agatgatgat gatgaacttg
 acgacgaggt ggaactgctg
 16861 cacgctaccg cggccaggcg acgggtacag tggaaaggtc
 gacgcgtaaa acgtgtttg
 16921 cgaccggca ccaccgtagt ctttacgccc ggtgagcgc
 ccaccgcac ctacaagcgc
 16981 gtgtatgatg aggtgtacgg cgacgaggac ctgcttgagc
 aggccaacga gcgcctcggg
 17041 gagtttgctt acggaagcg gcataaggac atgctggcgt
 tgccgctgga cgagggcaac
 17101 ccaacaccta gcctaaagcc cgtaacactg cagcagggtgc
 tgcccgcgct tgcaccgtcc
 17161 gaagaaaagc gcggcctaaa gcgcgagtct ggtgacttgg
 caccaccgt gcagctgatg
 17221 gtacccaagc gccagcgact ggaagatgtc ttggaaaaaa
 tgaccgtgga acctgggctg
 17281 gagcccaggg tccgcgtgcg gccaatcaag cagggtggcgc
 cgggactggg cgtgcagacc
 17341 gtggacgttc agataccac taccagtagc accagtattg
 ccaccgccac agagggcatg
 17401 gagacacaaa cgtccccggt tgccctcagcg gtggcggatg
 ccgcggtgca ggcggtcgc
 17461 gcggccgctg ccaagacctc tacggagggtg caaacggacc
 cgtggatggt tcgcgtttca
 17521 gcccccggc gcccgcgcgg ttcgaggaag tacggcgccg
 ccagcgcgct actgcccga
 17581 tatgcctac atccttccat tgcgcctacc cccggctatc
 gtggctacac ctaccgccc
 17641 agaagacgag caactaccgc acgcccgaacc accactggaa
 cccgcgcgcg ccgtcgcgct
 17701 cgcagcccgc tgctggcccc gatttccgtg cgcagggtgg
 ctgcggaagg aggcaggacc
 17761 ctggtgctgc caacagcgcg ctaccacccc agcatcgttt
 aaaagccggt ctttgtggtt

17821 cttgcagata tggccctcac ctgccgcctc cgtttcccgg
 tgccgggatt ccgaggaaga
 17881 atgcaccgta ggaggggcat ggccggccac ggcctgacgg
 gcggcatgcg tcgtgcgcac
 17941 caccggcggc ggcgcgcgtc gcaccgtcgc atgcgcggcg
 gtatcctgcc cctccttatt
 18001 ccaactgatcg ccgcgcgat tggcgccgtg cccggaattg
 catcogtggc cttgcaggcg
 18061 cagagacact gattaanaac aagttgcatg tggaaaaatc
 aaaataaaaa gtctggactc
 18121 tcacgctcgc ttggctctgt aactattttg tagaatggaa
 gacatcaact ttgcgtctct
 18181 ggccccgcga cacggctcgc gcccgttcat gggaaactgg
 caagatatcg gcaccagcaa
 18241 tatgagcggg ggcgccctca gctggggctc gctgtggagc
 ggcattaaaa atttcggttc
 18301 caccgttaag aactatggca gcaaggcctg gaacagcagc
 acaggccaga tgctgaggga
 18361 taagttgaaa gagcaaaatt tccaacaaaa ggtggtagat
 ggctggcct ctggcattag
 18421 cgggggtggtg gacctggcca accaggcagt gcaaaataag
 attaacagta agcttgatcc
 18481 ccgccctccc gtagaggagc ctccaccggc cgtggagaca
 gtgtctccag aggggctggtg
 18541 cgaaaagcgt ccgcgccccg acagggaaga aactctggtg
 acgcaaatag acgagcctcc
 18601 ctggtacgag gaggcactaa agcaaggcct gccaccacc
 cgteccatcg cgcccatggc
 18661 taccggagtg ctgggcccagc acacaccggt aacgctggac
 ctgcctcccc ccgcccagac
 18721 ccagcagaaa cctgtgctgc caggcccagc cgccgttgtt
 gtaaccgctc ctagccgcgc
 18781 gtccctgcgc ccgcgcccga gcggctccgc atcgttgccg
 cccgtagcca gtggcaactg
 18841 gcaaagcaca ctgaacagca tcgtgggtct gggggtgcaa
 tcctgaagc gccgacgatg
 18901 cttctgaata gctaacgtgt cgtatgtgtg tcatgtatgc
 gtccatgtcg ccgcccagag
 18961 agctgctgag ccgcccgcgc cccgctttcc aagatggcta
 ccccttcgat gatgccgag
 19021 tggctttaca tgcacatctc gggccaggac gcctcggagt
 acctgagccc cgggctggtg
 19081 cagtttgccc gcgccaccga gacgtacttc agcctgaata
 acaagtttag aaaccccacg
 19141 gtggcgcta cgcacgacgt gaccacagac cggctcccagc
 gtttgacgct gcggttcac
 19201 cctgtggacc gtgaggatac tgcgtactcg tacaaggcgc
 ggttcaccct agctgtgggt
 19261 gataaccgtg tgctggacat ggcttcacg tactttgaca
 tccgcccgt gctggacagg

19321 ggccctactt ttaagcccta ctctggcact gcctacaacg
 ccctggctcc caagggtgcc
 19381 ccaaatcctt gcgaatggga tgaagctgct actgctcttg
 aaataaacct agaagaagag
 19441 gacgatgaca acgaagacga agtagacgag caagctgagc
 agcaaaaaac tcacgtatth
 19501 gggcaggcgc cttattctgg tataaatatt acaaaggagg
 gtattcaaht aggtgtcga
 19561 ggtcaaacac ctaahtatgc cgataaaaca tttcaacctg
 aacctcaaht aggagaatct
 19621 cagtggtagc aaactgaaht taatcatgca gctgggagag
 tccttaaaaa gactacccca
 19681 atgaaacctat gttacgggtc atatgcaaaa cccacaaatg
 aaaatggagg gcaaggcatt
 19741 cttgtaaagc aacaaaatgg aaagctagaa agtcaagtgg
 aatgcaatth tttctcaact
 19801 actgaggcga ccgcaggcaa tgggtgataac ttgactccta
 aagtggatth gtacagtga
 19861 gatgtagata tagaaacccc agacactcat atthcttaca
 tgcccactat taaggaaggt
 19921 aactcacgag aactaatggg ccaacaatct atgcccaca
 ggccataatth cattgctthth
 19981 agggacaatth ttattgggtct aatgtattac aacagcacgg
 gtaaatatggg tgttctggcg
 20041 ggccaagcat cgcagttgaa tgctgttgta gatttgcaag
 acagaaacac agagctthca
 20101 taccagctth tgcttgattc cattgggtgat agaaccaggt
 actththctat gtggaatcag
 20161 gctgttgaca gctatgatcc agatgttaga attattgaaa
 atcatggaac tgaagatgaa
 20221 cttccaaatth actgctthcc actgggaggt gtgattaata
 cagagactct taccaaggta
 20281 aaacctaaaa caggtcagga aaatggatgg gaaaaagatg
 ctacagaatth ttcagataaa
 20341 aatgaaataa gagttggaaa taatthtgcc atggaaatca
 atctaaatgc caacctgtgg
 20401 agaaatthcc tgtactcaa catagcgctg tatttgcccg
 acaagctaaa gtacagtcct
 20461 tccaacgtaa aaatthctga taaccacaaac acctacgact
 acatgaacaa gcgagtggtg
 20521 gctcccgggt tagtggactg ctacattaac cttggagcac
 gctggtcctt tgactatatg
 20581 gacaacgtca accatthta ccaccaccgc aatgctggcc
 tgcgctaccg ctcaatgttg
 20641 ctgggcaatg gtcgctatgt gccctthcac atccaggtgc
 ctcagaagtt cthtgccatt
 20701 aaaaacctcc thctctgccc gggctcatic acctacgagt
 ggaactthcag gaaggatgtt
 20761 aacatgggtc tgcagagctc cctaggaaht gacttaaggg
 ttgacggagc cagcattaag

20821 ttgatagca ttgccttta cgccaccttc ttccccatgg
 cccacaacac cgctccacg
 20881 cttgaggcca tgcttagaaa cgacaccaac gaccagtcct
 ttaacgacta tctctccgcc
 20941 gccaacatgc tctaccctat acccgccaac gctaccaacg
 tgcccatatc catcccctcc
 21001 cgcaactggg cggttttccg cggctgggccc ttcacgcgcc
 ttaagactaa ggaaacccca
 21061 tcaactgggct cgggctacga cccttattac acctactctg
 gctctatacc ctacctagat
 21121 ggaacctttt acctcaocca cacctttaag aagggtggcca
 ttacctttga ctcttctgtc
 21181 agctggcctg gcaatgaccg cctgcttacc cccaacgagt
 ttgaaattaa gcgctcagtt
 21241 gacggggagg gttacaacgt tgcccagtggt aacatgacca
 aagactgggt cctggtacaa
 21301 atgctagcta actacaacat tggctaccag ggcttctata
 toccagagag ctacaaggac
 21361 cgcatgtact ccttctttag aaacttccag cccatgagcc
 gtcaggtggg gatgatact
 21421 aaatacaagg actaccaaca ggtgggcatc ctacaccaac
 acaacaactc tggatttgtt
 21481 ggctaccttg cccccacct gcgcgaagga caggcctacc
 ctgctaactt cccctatccg
 21541 cttataggca agaccgcagt tgacagcatt acccagaaaa
 agtttctttg cgatcgcacc
 21601 ctttggcgca tcccattctc cagtaacttt atgtccatgg
 gcgcactcac agacctgggc
 21661 caaaccttc tctacgcca ctccgcccac gcgctagaca
 tgacttttga ggtggatccc
 21721 atggacgagc ccacccttct ttatgttttg tttgaagtct
 ttgacgtggg ccgtgtgcac
 21781 cggccgcacc gcggcgctcat cgaaaccgtg tacctgcgca
 cgcccttctc ggccggcaac
 21841 gccacaacat aaagaagcaa gcaacatcaa caacagctgc
 cgccatgggc tccagtgagc
 21901 aggaactgaa agccattgtc aaagatcttg gttgtgggccc
 atattttttg ggcacctatg
 21961 acaagecgtt tccaggcttt gtttctccac acaagctcgc
 ctgcgccata gtcaatacgg
 22021 ccggtcgcga gactgggggc gtacactgga tggcctttgc
 ctggaaccgg cactcaaaaa
 22081 catgctacct ctttgagccc tttggctttt ctgaccagcg
 actcaagcag gtttaccagt
 22141 ttgagtacga gtcactcctg cgccgtagcg ccattgcttc
 ttcccccgac cgctgtataa
 22201 cgctggaaaa gtccacccaa agcgtacagg ggcccactc
 ggccgcctgt ggactattct
 22261 gctgcatggt tctccacgcc tttgccaaact ggccccaaac
 tcccatggat cacaacccca

22321 ccatgaacct tattaccggg gtaccaact ccatgctcaa
 cagtccccag gtacagccca
 22381 ccctgcgtcg caaccaggaa cagctctaca gcttcttggg
 gcgccactcg ccctacttcc
 22441 gcagccacag tgcgcagatt aggagcgcca cttctttttg
 tcacttgaaa aacatgtaaa
 22501 aataatgtac tagagacact ttcaataaag gcaaatgctt
 ttatttgtac actctcgggt
 22561 gattatttac cccaccctt gccgtctgcg ccgtttaaaa
 atcaaagggg ttctgcccgc
 22621 catcgctatg cgccactggc agggacacgt tgcgatactg
 gtgttttagtg ctccacttaa
 22681 actcaggcac aaccatccgc ggcagctcgg tgaagttttc
 actccacagg ctgcgcacca
 22741 tcaccaacgc gtttagcagg tcgggcgccg atatcttgaa
 gtcgcagttg gggcctccgc
 22801 cctgcgcgcg cgagttgcga tacacagggt tgcagcactg
 gaacactatc agcgcgggt
 22861 ggtgcacgct ggccagcacg ctcttgcgg agatcagatc
 cgcgtccagg tctcgcgct
 22921 tgetcagggc gaacggagtc aactttggta gctgccttcc
 caaaaagggc gcgtgccag
 22981 gctttgagtt gcactcgcac cgtagtggca tcaaaagggt
 accgtgcccg gtctgggctg
 23041 taggatacag cgcctgcata aaagccttga tctgcttaaa
 agccacctga gcctttgcgc
 23101 cttcagagaa gaacatgccg caagacttgc cggaaaactg
 attggccgga caggccgctg
 23161 cgtgcacgca gcaccttgcg tcgggtgttg agatctgcac
 cacatttcgg ccccaccggt
 23221 tcttcacgat cttggccttg ctagactgct ccttcagcgc
 gcgctgcccg ttttcgctcg
 23281 tcacatccat ttcaatcacg tgctccttat ttatcataat
 gcttccgtgt agacacttaa
 23341 gctcgccttc gatctcagcg cagcgggtgca gccacaacgc
 gcagcccgtg ggctcgtgat
 23401 gcttgtaggt cacctctgca aacgactgca ggtacgcctg
 caggaatcgc cccatcatcg
 23461 tcacaaaggt cttggtgctg gtgaagggtca gctgcaaccc
 gcggtgctcc tcgttcagcc
 23521 aggtcttgca tacggccgcc agagcttcca cttggtcagg
 cagtgtttg aagttcgcct
 23581 ttagatcgtt atccacgtgg tacttgtcca tcagcgcgcg
 cgcagcctcc atgcccttct
 23641 cccacgcaga cacgatcggc aactcagcg ggttcatcac
 cgtaatttca ctttccgctt
 23701 cgctgggctc ttctcttcc tcttgcgtcc gcataccacg
 cgccactggg tcgtcttcat
 23761 tcagccgccc cactgtgcgc ttacctcctt tgccatgctt
 gattagcacc ggtgggttgc

23821 tgaaacccac catttgtagc gccacatctt ctctttcttc
 ctcgctgtcc acgattacct
 23881 ctggtgatgg cgggcgctcg ggcttgggag aagggcgctt
 ctttttcttc ttgggcgcaa
 23941 tggccaaatc cgccgccgag gtcgatggcc gcgggctggg
 tgtgcgcggc accagcgcgt
 24001 cttgtgatga gtcttctcgc tcctcggact cgatacgcgc
 cctcatccgc ttttttgggg
 24061 gcgcccgggg aggcggcggc gacggggacg gggacgacac
 gtcctccatg gttgggggac
 24121 gtcgcgccgc accgcgtccg cgctcggggg tggtttcgcg
 ctgctcctct tcccgactgg
 24181 ccatttctct ctctatagg cagaaaaaga tcatggagtc
 agtcgagaag aaggacagcc
 24241 taaccgcccc ctctgagttc gccaccaccg cctccaccga
 tgccgccaac gcgcctacca
 24301 ctttccccgt cgaggcacc cgccttgagg aggaggaagt
 gattatcgag caggaccag
 24361 gttttgtaag cgaagacgac gaggaccgct cagtaccaac
 agaggataaa aagcaagacc
 24421 aggacaacgc agaggcaaac gaggaacaag tcgggcgggg
 ggacgaaagg catggcgact
 24481 acctagatgt gggagacgac gtgctgttga agcatctgca
 gcgccagtgc gccattatct
 24541 gcgacgcgtt gcaagagcgc agcgatgtgc ccctcgccat
 agcggatgtc agccttgctt
 24601 acgaacgcca cctattctca ccgcgcgtac ccccaaacg
 ccaagaaaac ggcacatgcg
 24661 agcccaacc gcgcctcaac ttctaccccg tatttgccgt
 gccagaggtg cttgccacct
 24721 atcacatctt tttccaaaac tgcaagatac ccctatcctg
 ccgtgccaac cgcagccgag
 24781 cggacaagca gctggccttg cggcagggcg ctgtcatacc
 tgatatcgcc tcgctcaacg
 24841 aagtgccaaa aatctttgag ggtcttggac gcgacgagaa
 gcgcgcggca aacgctctgc
 24901 aacaggaaaa cagcgaaaat gaaagtcact ctggagtgtt
 ggtggaactc gagggtgaca
 24961 acgcgcgcct agccgtacta aaacgcagca tcgaggtcac
 ccactttgcc taccggcac
 25021 ttaacctacc cccaaggtc atgagcacag tcatgagtga
 gctgatcgtg cgccgtgccc
 25081 agcccctgga gagggatgca aatttgcaag aacaaacaga
 ggagggccta cccgcagttg
 25141 gcgacgagca gctagcgcgc tggcttcaaa cgcgcgagcc
 tgccgacttg gaggagcgc
 25201 gcaaactaat gatggccgca gtgctcgta ccgtggagct
 tgagtgcattg cagcggttct
 25261 ttgctgacct ggagatgcag' cgcaagctag aggaaacatt
 gcactacacc tttcgacagg

25321 gctacgtacg ccaggcctgc aagatctcca acgtggagct
 ctgcaacctg gtctcctacc
 25381 ttggaatttt gcacgaaaac cgccttgggc aaaacgtgct
 tcattccacg ctcaagggcg
 25441 aggcgcgccg cgactacgtc cgcgactgcg tttacttatt
 tctatgctac acctggcaga
 25501 cggccatggg cgtttggcag cagtgcttgg aggagtgcaa
 cctcaaggag ctgcagaaac
 25561 tgctaaagca aaacttgaag gacctatgga cggccttcaa
 cgagcgctcc gtggccgcg
 25621 acctggcgga catcattttc cccgaacgcc tgcttaaaac
 cctgcaacag ggtctgccag
 25681 acttcaccag tcaaagcatg ttgcagaact ttaggaactt
 tatcctagag cgctcaggaa
 25741 tcttgcccgc cacctgctgt gcacttccta gcgactttgt
 gccattaag taccgcaat
 25801 gccctccgcc gctttggggc cactgctacc ttctgcagct
 agccaactac ctgacctacc
 25861 actctgacat aatggaagac gtgagcggtg acggtctact
 ggagtgtcac tgtcgtgca
 25921 acctatgcac cccgcaccgc tccctggttt gcaattcgca
 gctgcttaac gaaagtcaaa
 25981 ttatcggtac ctttgagctg cagggtccct cgctgacga
 aaagtccgcg gctccgggg
 26041 tgaaactcac tccggggctg tggacgtcgg cttaccttcg
 caaatttcta cctgaggact
 26101 accacgcca cgagattagg ttctacgaag accaatcccg
 cccgccaat gcggagctta
 26161 ccgctgcgt cattaccag gccacattc ttggccaatt
 gcaagccatc acaaagccc
 26221 gccaaagatt tctgctacga aaggacggg gggtttactt
 ggacccccag tccggcgagg
 26281 agctcaacc aatcccccg ccgccgcagc cctatcagca
 gcagccgcg gcccttgctt
 26341 cccaggatgg cacccaaaa gaagctgcag ctgcccgcgc
 caccacgga cgaggaggaa
 26401 tactgggaca gtcaggcaga ggaggtttg gacgaggagg
 aggaggacat gatggaagac
 26461 tgggagagcc tagacgagga agcttccgag gtcgaagagg
 tgtcagacga aacaccgtca
 26521 cctcgggtcg cattcccctc gccggcgccc cagaaatcgg
 caaccggttc cagcatgget
 26581 acaacctccg ctctcaggg gccgcccgca ctgcccgttc
 gccgaccaa ccgtagatgg
 26641 gacaccactg gaaccagggc cggttaagtcc aagcagccgc
 cgccgtagc ccaagagcaa
 26701 caacagcgc aaggctaccg ctcatggcgc gggcacaaga
 acgcatagt tgcttgcttg
 26761 caagactgtg ggggcaacat ctcttcgcc cgccgcttc
 ttctctacca tcacggcgtg

26821 gccttccccc gtaacatcct gcattactac cgtcatctct
 acagcccata ctgcaccggc
 26881 ggcagcggca gggcagcaa cagcagcggc cacacagaag
 caaaggcgac cggatagcaa
 26941 gactctgaca aagcccaaga aatccacagc ggcggcagca
 gcaggaggag gagcgctgcg
 27001 tctggcgccc aacgaaccgg tatcgaccgg cgagcttaga
 aacaggattt ttcccactct
 27061 gtatgctata tttcaacaga gcagggggcca agaacaagag
 ctgaaaataa aaaacaggtc
 27121 tctgcgatcc ctcaccggca gctgctgta tcacaaaagc
 gaagatcagc ttcggcgcac
 27181 gctggaagac gcgagggtc tcttcagtaa atactgcgcg
 ctgactctta aggactagtt
 27241 tcgcgccctt tctcaaattt aagcgcgaaa actacgtcat
 ctccagcggc cacaccggc
 27301 gccagcacct gtcgtcagcg ccattatgag caaggaaatt
 cccacgccct acatgtggag
 27361 ttaccagcca caaatgggac ttgcggtgag agctgcccac
 gactactcaa cccgaataaa
 27421 ctacatgagc gcgggacccc acatgatata ccgggtcaac
 ggaatccgcg cccaccgaaa
 27481 ccgaattctc ttggaacagg cggctattac caccacacct
 cgtaataacc ttaatccccg
 27541 tagttggccc gctgccctgg tgtaccagga aagtcccgct
 cccaccactg tggacttcc
 27601 cagagacgcc caggccgaag ttcagatgac taactcaggg
 gcgagcttg cgggcggcct
 27661 tcgtcacagg gtgcggtcgc ccgggcaggg tataactcac
 ctgacaatca gagggcgagg
 27721 tattcagctc aacgacgagt cggtgagctc ctcgcttggg
 ctccgtccgg acgggacatt
 27781 tcagatcggc ggcgccggcc gtccttcatt cacgcctcgt
 caggcaatcc taactctgca
 27841 gacctcgtcc tctgagccgc gctctggagg cattggaact
 ctgcaattta ttgaggagtt
 27901 tgtgccatcg gtctacttta accccttctc gggacctccc
 ggccactatc cggatcaatt
 27961 tattcctaac tttgacgcgg taaaggactc ggcggacggc
 tacgactgaa tgtaagtgg
 28021 agaggcagag caactgcgcc tgaaacacct ggtccactgt
 cgccgccaca agtgctttgc
 28081 ccgcgactcc ggtgagtttt gctactttga attgcccagag
 gatcatatcg agggcccggc
 28141 gcacggcgtc cggcttaccg ccagggaga gcttgcccgt
 agcctgattc gggagtttac
 28201 ccagcgcgcc ctgctagttg agcgggacag gggaccctgt
 gttctcactg tgatttgcaa
 28261 ctgtcctaac cttggattac atcaagatcc tctagttata
 actagagtagc ccggggatct

28321 tattcccttt aactaataaa aaaaaataat aaagcatcac
 ttacttaaaa tcagtttagca
 28381 aatttctgtc cagtttattc agcagcacct ccttgcctc
 ctcccagctc tggatttgca
 28441 gcttcctcct ggctgcaaac tttctccaca atctaaatgg
 aatgtcagtt tcctcctggt
 28501 cctgtccatc cgcacccact atcttcatgt tgttgcagat
 gaagcgcgca agaccgtctg
 28561 aagatacctt caaccccgtg tatccatgat acacggaaac
 cggtcctcca actgtgcctt
 28621 ttcttactcc tccctttgta tcccccaatg ggtttcaaga
 gagtccccct ggggtactct
 28681 ctttgcgctt atccgaacct ctagttacct ccaatggcat
 gcttgcgctc aaaatgggca
 28741 acggcctctc tctggacgag gccggcaacc ttacctcca
 aatgtaacc actgtgagcc
 28801 cacctctcaa aaaaaccaag tcaaacataa acctggaat
 atctgcacc ctcacagtta
 28861 cctcagaagc ctaactgtg gctgcccggc cacctcta
 ggtcgcccggc aacacactca
 28921 ccatgcaatc acaggccccg ctaaccgtgc acgactcca
 acttagcatt gccacccaag
 28981 gaccctcac agtgtcagaa ggaaagctag ccctgcaaac
 atcaggcccc ctcaccacca
 29041 ccgatagcag tacccttact atcactgcct cccccctct
 aactactgcc actggtagct
 29101 tgggcattga cttgaaagag cccatttata cacaaaatgg
 aaaactagga ctaaagtacg
 29161 gggctccttt gcatgtaaca gacgacctaa acactttgac
 cgtagcaact ggtccaggtg
 29221 tgactattaa taatacttcc ttgcaaacta aagttactgg
 agccttgggt tttgattcac
 29281 aaggcaatat gcaacttaat gtagcaggag gactaaggat
 tgattctcaa aacagacgcc
 29341 ttatacttga tgttagttat ccgtttgatg ctcaaaaacca
 actaaatcta agactaggac
 29401 agggccctct ttttataaac tcagcccaca acttggatat
 taactacaac aaaggccttt
 29461 acttgtttac agcttcaaac aattccaaaa agcttgaggt
 taacctaaagc actgccaagg
 29521 ggttgatggt tgacgctaca gccatagcca ttaatgcagg
 agatgggctt gaatttggtt
 29581 cacctaattg accaaacaca aatcccctca aaacaaaaat
 tggccatggc ctagaatttg
 29641 attcaaaaaa ggctatgggt cctaaactag gaactggcct
 tagttttgac agcacaggtg
 29701 ccattacagt aggaaacaaa aataatgata agctaacttt
 gtggaccaca ccagctccat
 29761 ctctaactg tagactaaat gcagagaaag atgctaaact
 cactttggtc ttaacaaaat

29821 gtggcagtc aatacttgct acagtttcag ttttggctgt
 taaaggcagt ttggctccaa
 29881 tatctggaac agttcaaagt gctcatctta ttataagatt
 tgacgaaaat ggagtgctac
 29941 taaacaattc cttcctggac ccagaatatt ggaactttag
 aatggagat cttactgaag
 30001 gcacagccta taaaaacgct gttggattta tgccctaacct
 atcagcttat ccaaaatctc
 30061 acggtaaaac tgccaaaagt aacattgtca gtcaagttta
 cttaaacgga gacaaaacta
 30121 aacctgtaac actaaccatt aactaaacg gtacacagga
 aacaggagac acaactccaa
 30181 gtgcatactc tatgtcattt tcatgggact ggtctggcca
 caactacatt aatgaaatat
 30241 ttgccacatc ctcttact tttcataca ttgcccaaga
 ataaagaatc gtttgtgta
 30301 tgtttcaacg tgtttatatt tcaattgcag aaaatttcaa
 gtcatttttc attcagtagt
 30361 atagccccac caccacatag cttatacaga tcaccgtacc
 ttaatcaaac tcacagaacc
 30421 ctagtattca acctgccacc tccctcccaa cacacagagt
 acacagtcct ttctccccgg
 30481 ctggccttaa aaagcatcat atcatgggta acagacatat
 tcttaggtgt tatattccac
 30541 acggtttcct gtcgagccaa acgctcatca gtgatattaa
 taaactcccc gggcagctca
 30601 cttaagttca tgctgctgct cagctgctga gccacaggct
 gctgtccaac ttgcggttgc
 30661 ttaacgggcg gcgaaggaga agtccacgcc tacatggggg
 tagagtcata atcgtgcatc
 30721 aggatagggc ggtggtgctg cagcagcgcg cgaataaact
 gctgccgccc cgctccgctc
 30781 ctgcaggaat acaacatggc agtggctctcc tcagcgatga
 ttgcaccgc ccgcagcata
 30841 aggcgccttg tcctccgggc acagcagcgc acctgatct
 cacttaaate agcacagtaa
 30901 ctgcagcaca gcaccacaat attggtcaaa atcccacagt
 gcaaggecgt gtatccaaag
 30961 ctcatggcgg ggaccacaga acccagctgg ccatcatacc
 acaagcgcag gtagattaag
 31021 tggcgacccc tcataaacac gctggacata aacattacct
 cttttggcat gttgtaattc
 31081 accacctccc ggtaccatat aaacctctga ttaaakatgg
 cgccatccac caccatccta
 31141 aaccagctgg ccaaaacctg cccgccggct atactactgca
 ggaaccggg actggaacaa
 31201 tgacagtga gagcccagga ctcgtaacca tggatcatca
 tgctcgtcat gatatcaatg
 31261 ttggcacaac acaggcacac gtgcatacac ttctcagga
 ttacaagctc ctccccgctt

31321 agaaccatat cccagggaac aaccattcc tgaatcagcg
 taaatcccac actgcaggga
 31381 agacctcgca cgtaactcac gttgtgcatt gtcaaagtgt
 tacattcggg cagcagcgga
 31441 tgatcctcca gtaggtagc gcgggtttct gtctcaaaag
 gaggtagacg atccctactg
 31501 tacggagtgc gccgagacaa ccgagatcgt gttggtcgta
 gtgtcatgcc aaatggaacg
 31561 ccggacgtag tcatatttcc tgaagcaaaa ccaggtgcgg
 gcgtgacaaa cagatctgog
 31621 tctccggtct cgccgcttag atcgctctgt gtagtagttg
 tagtatatcc actctctcaa
 31681 agcatccagg cgccccctgg ctccgggttc tatgtaaact
 ccttcatgog ccgctgcct
 31741 gataacatcc accaccgag aataagccac acccagccaa
 cctacacatt cgttctgca
 31801 gtcacacacg ggaggagcgg gaagagctgg aagaaccatg
 tttttttttt tattccaaaa
 31861 gattatccaa aacctcaaaa tgaagatcta ttaagtgaac
 gcgctcccct ccggtggcgt
 31921 ggtcaaactc tacagccaaa gaacagataa tggcatttgt
 aagatgttgc acaatggctt
 31981 ccaaaaggca aacggccctc acgtccaagt ggacgtaaag
 gctaaaccct tcagggtgaa
 32041 tctcctctat aaacattcca gcaccttcaa ccatgcccua
 ataattctca tctcggccacc
 32101 ttctcaatat atctctaagc aaatcccga tattaagtcc
 ggccattgta aaaatctgct
 32161 ccagagcgcc ctccacctc agcctcaagc agcgaatcat
 gattgcaaaa attcaggttc
 32221 ctacagacc tgtataagat tcaaaagcgg aacattaaca
 aaaataccgc gatcccgtag
 32281 gtcccttcgc agggccagct gaacataatc gtgcaggtct
 gcacggacca gcgoggccac
 32341 ttccccgcca ggaacctga caaaagaacc cacactgatt
 atgacacgca tactcggagc
 32401 tatgctaacc agcgtagccc cgatgtaagc tttgttgcatt
 gggcggcgat ataaaatgca
 32461 aggtgctgct caaaaaatca ggcaaagcct cgcgcaaaaa
 agaaagcaca tcgtagtcat
 32521 gctcatgcag ataaaggcag gtaagctccg gaaccaccac
 agaaaaagac accatttttc
 32581 tctcaaacat gtctgogggt ttctgcataa acacaaaata
 aaatacaaaa aaaacattta
 32641 aacattagaa gcctgtctta caacaggaaa aacaaccott
 ataagcataa gacggactac
 32701 ggccatgccg gcgtgaccgt aaaaaaactg gtcaccgtga
 ttaaaaagca ccaccgacag
 32761 ctctcgggtc atgtccggag tcataatgta agactcggta
 aacacatcag gttgattcat

32821 cggtcagtgc taaaaagcga ccgaaatagc ccgggggaat
 acatacccgc aggcgtagag
 32881 acaacattac agccccata ggaggtataa caaaattaat
 aggagagaaa aacacataaa
 32941 cacctgaaaa accctcctgc ctaggcaaaa tagcacctc
 ccgctccaga acaacataca
 33001 gcgcttcaca gcggcagcct aacagtcagc cttaccagta
 aaaaagaaaa cctattaata
 33061 aacaccact cgacacggca ccagctcaat cagtcacagt
 gtaaaaaagg gccaaagtga
 33121 gagcgagtat atataggact aaaaaatgac gtaacgggta
 aagtccacaa aaaacaccca
 33181 gaaaaccgca cgcgaaaccta cgcccagaaa cgaaagccaa
 aaaaccaca acttcctcaa
 33241 atcgtcactt ccgttttccc acgttacgta acttcccatt
 ttaagaaaac tacaattccc
 33301 aacacataca agttactcgg ccctaaaacc tacgtcacc
 gccccgttcc cacgccccgc
 33361 gccacgtcac aaactccacc ccctcattat catattggct
 tcaatccaaa ataaggtata
 33421 ttattgatga tnnnnnttaa t

REIVINDICACIONES

1. Un vector de expresión adenoviral quimérico, comprendiendo dicho vector un casete de expresión que comprende los siguientes elementos:

- (a) un primer promotor unido operativamente a un ácido nucleico que codifica un agonista de receptor 3 tipo Toll (TLR-3), en el que el agonista de TLR-3 es ARNbc heterólogo y en el que el ácido nucleico que codifica el agonista de TLR-3 comprende una secuencia seleccionada del grupo que consiste en: las SEQ ID NO: 3, 4, 5, 8, 9, 10, 11 y 12; y
- (b) un segundo promotor unido operativamente a un ácido nucleico que codifica un polipéptido heterólogo inmunogénico.

2. El vector de expresión adenoviral quimérico de la reivindicación 1, en el que el primer promotor y el segundo promotor son los mismos, opcionalmente en el que el primer promotor y el segundo promotor son, cada uno, un promotor del CMV.

3. Una composición inmunogénica que comprende el vector de expresión de una cualquiera de las reivindicaciones precedentes y un vehículo farmacéuticamente aceptable.

4. Un vector de expresión adenoviral quimérico, comprendiendo dicho vector un casete de expresión que comprende los siguientes elementos:

- (a) un primer promotor unido operativamente a un ácido nucleico que codifica un agonista de receptor 3 tipo Toll (TLR-3), en el que el agonista de TLR-3 es ARNbc heterólogo; y
- (b) un segundo promotor unido operativamente a un ácido nucleico que codifica un polipéptido heterólogo inmunogénico

para su uso en un método de tratamiento que provoca una respuesta inmunitaria, comprendiendo dicho método administrar una cantidad inmunogénicamente eficaz del vector a un sujeto mamífero, tal como un ser humano, en el que la respuesta inmunitaria se dirige contra el polipéptido heterólogo, y en el que la vía de administración se selecciona del grupo que consiste en oral, intranasal y mucosa, tal como vaginal.

5. El vector para el uso de la reivindicación 4, en el que el polipéptido heterólogo se expresa en una célula seleccionada del grupo que consiste en una célula dendrítica, una célula de micropliegue y una célula epitelial intestinal.

6. El vector de la reivindicación 1, en el que el ARNbc heterólogo es ARN de horquilla corta (ARNhc).

7. Una composición inmunogénica, comprendiendo dicha composición:

- (a) un vector de expresión adenoviral quimérico que comprende un promotor unido operativamente a un ácido nucleico que codifica un polipéptido heterólogo inmunogénico;
- (b) un agonista de TLR-3, en el que el agonista de TLR-3 es ARNbc heterólogo; y
- (c) un vehículo farmacéuticamente aceptable;

para su uso en un método de tratamiento que provoca una respuesta inmunitaria, comprendiendo dicho método administrar una cantidad inmunogénicamente eficaz de la composición a un sujeto mamífero, tal como un ser humano, en el que la respuesta inmunitaria se dirige contra el polipéptido heterólogo, y en la que la vía de administración se selecciona del grupo que consiste en oral, intranasal y mucosa, tal como vaginal.

8. La composición para el uso de la reivindicación 7, en la que un promotor es un promotor del CMV.

9. La composición para el uso de la reivindicación 7, en la que el agonista de TLR-3 es un ácido nucleico que comprende una secuencia seleccionada del grupo que consiste en las SEQ ID NO: 3, 4, 5, 8, 9, 10, 11 y 12.

10. La composición para el uso de la reivindicación 7, en la que el polipéptido heterólogo se expresa en una célula seleccionada del grupo que consiste en una célula dendrítica, una célula de micropliegue y una célula epitelial intestinal.

11. La composición para el uso de la reivindicación 7, en la que el ARNbc heterólogo es ARN de horquilla corta (ARNhc).

12. El vector de la reivindicación 1, el vector para el uso de la reivindicación 4, o la composición para el uso de la reivindicación 7, en el que el polipéptido heterólogo es:

- (a) un polipéptido env del VIH, tal como un polipéptido seleccionado del grupo que consiste en gp41, gp120 y

gp160;

(b) un polipéptido HA de la gripe;

(c) un antígeno del virus del herpes simple (VHS); o

(d) un antígeno del virus del papiloma humano (VPH).

5

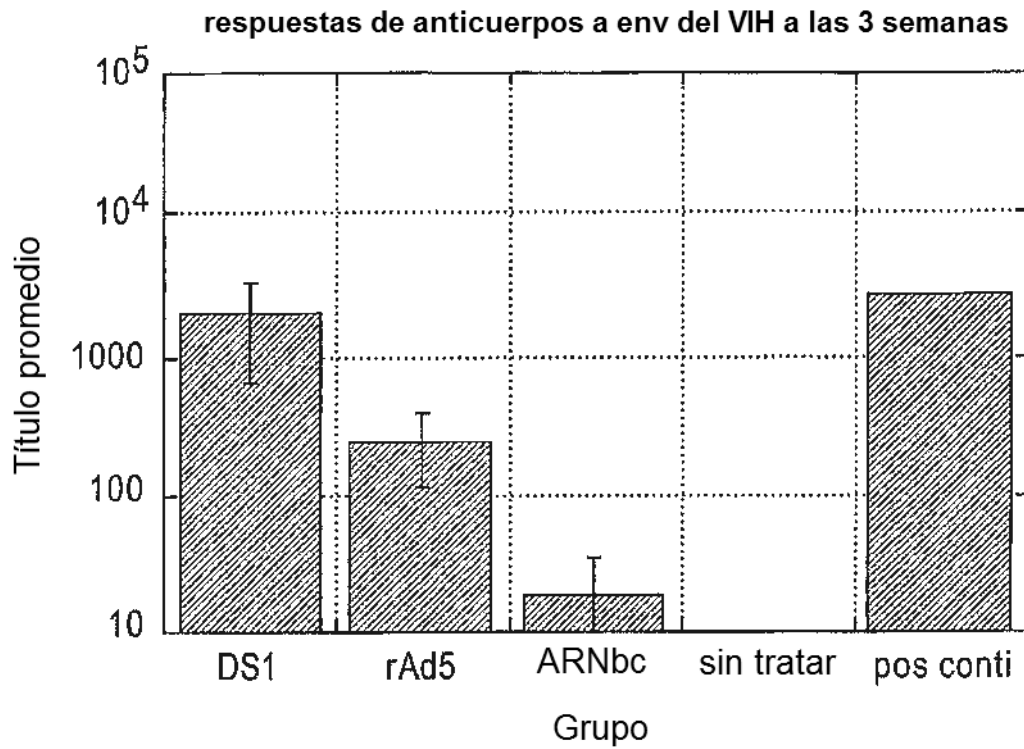


FIG. 1A

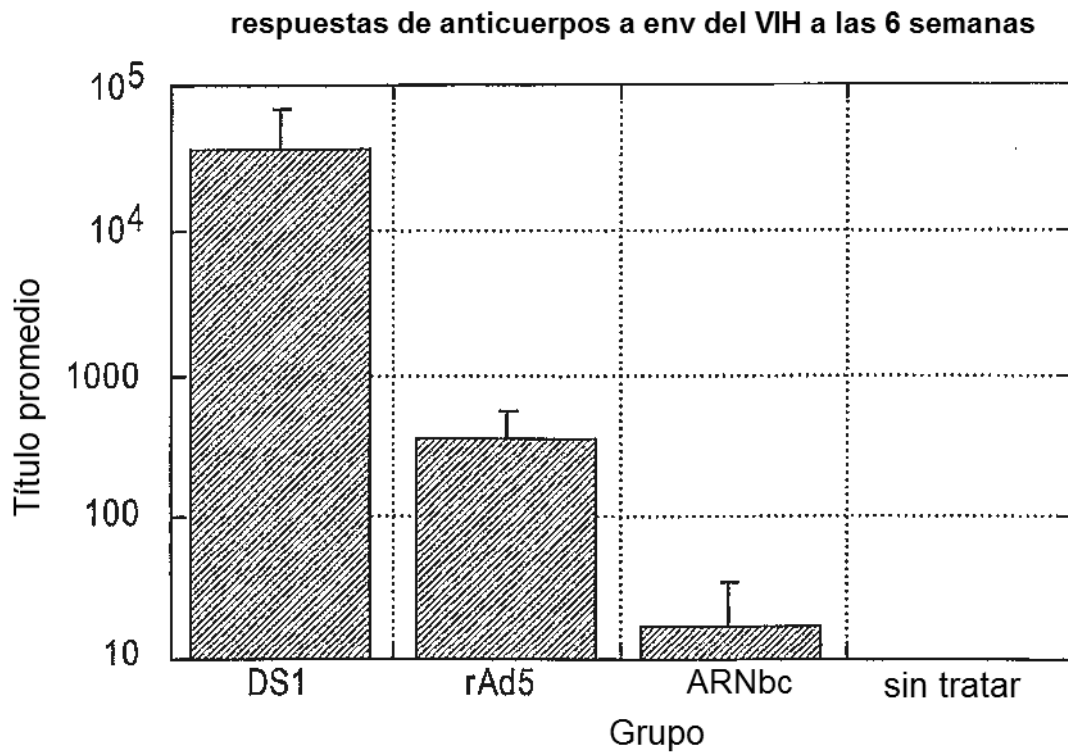


FIG. 1B

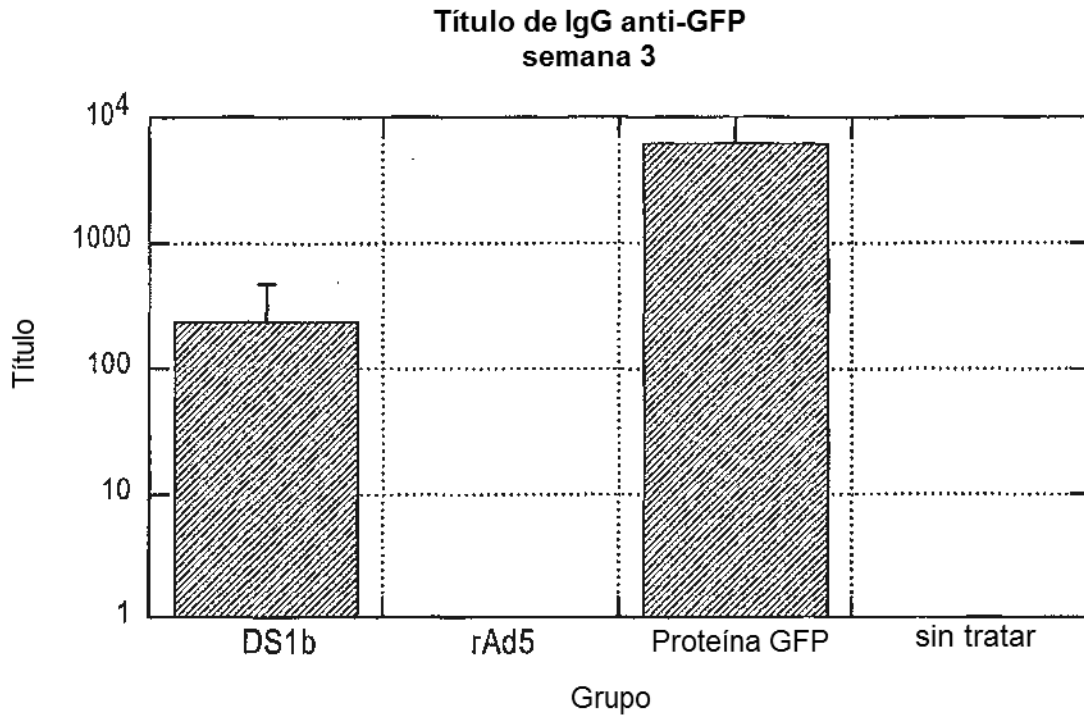


FIG. 2A

**% de tetrámero+ (péptido de GFP) de los linfocitos T CD8+
tras la vacunación con Ad-GFP**

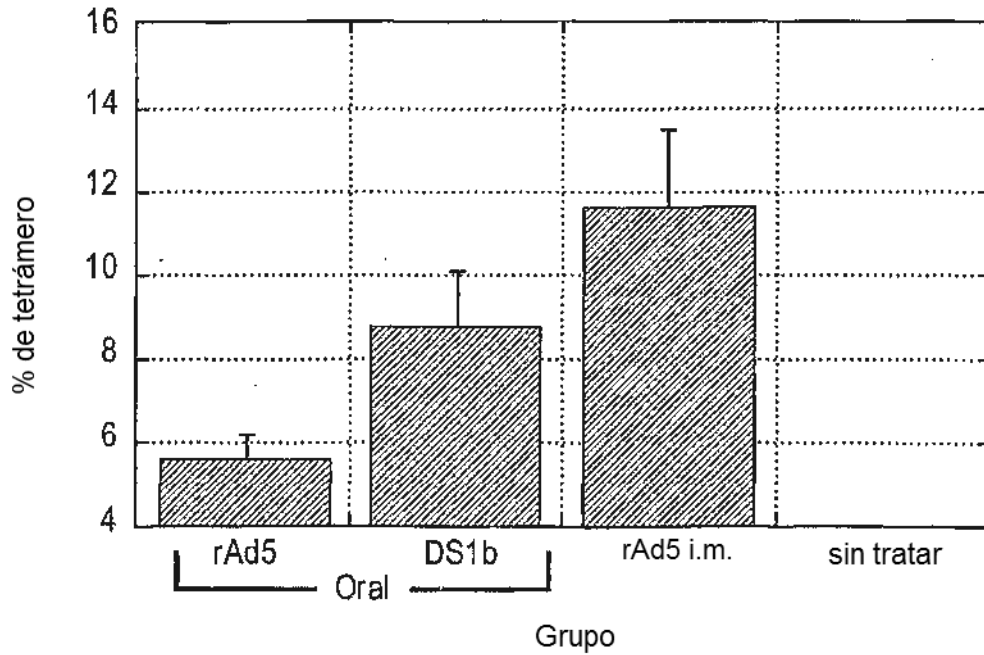


FIG. 2B

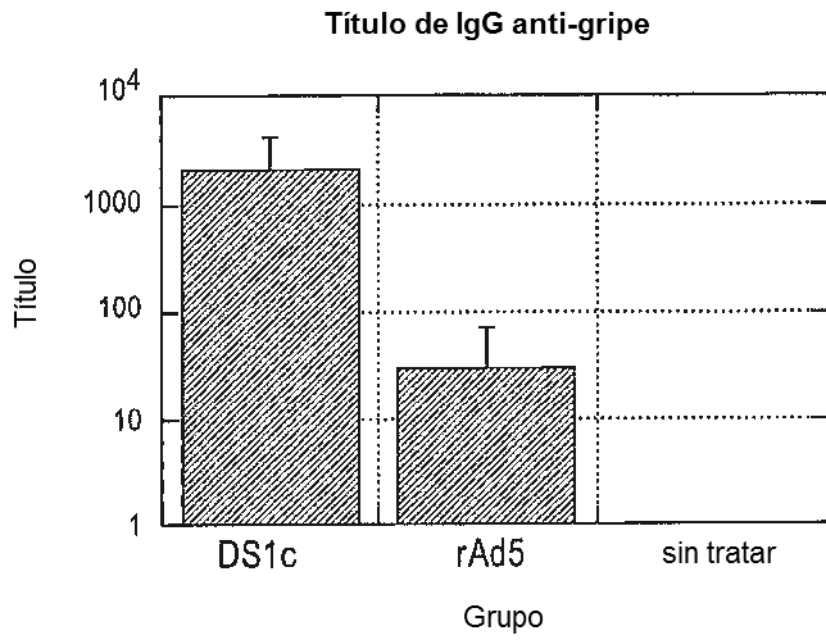


FIG. 2C

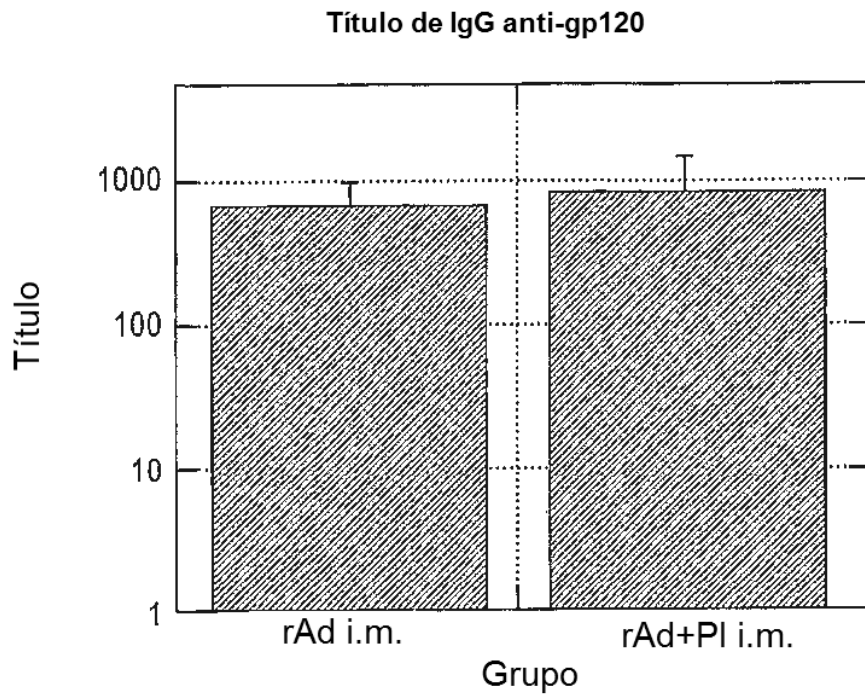


FIG. 3A

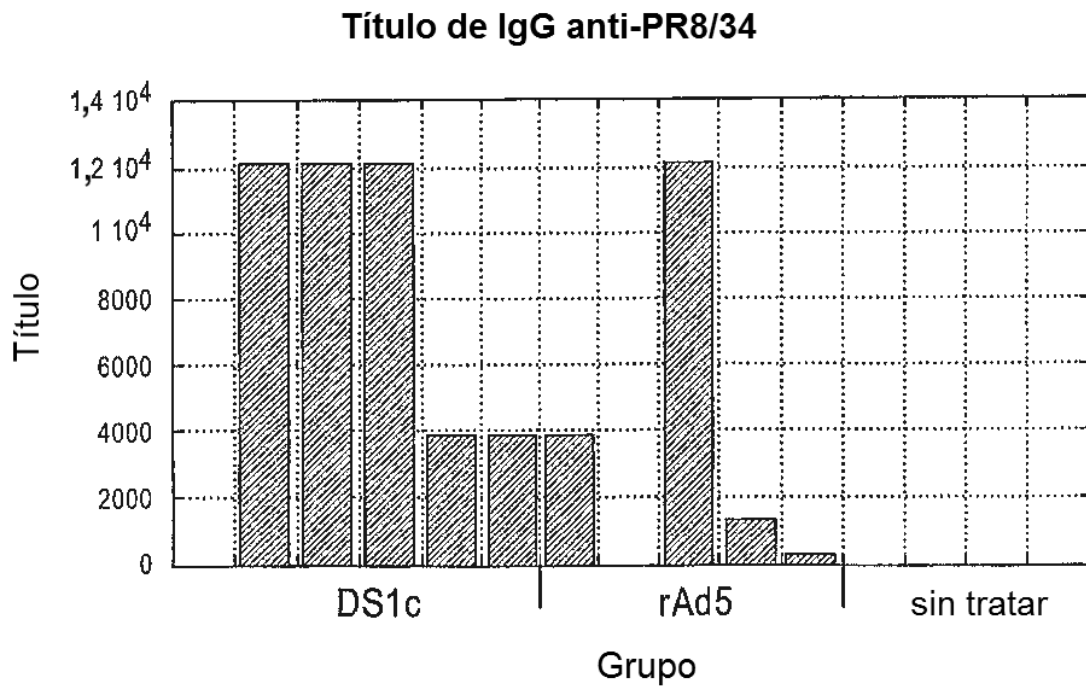


FIG. 3B

Secreción de células dendríticas de IL-6
en respuesta a estímulo

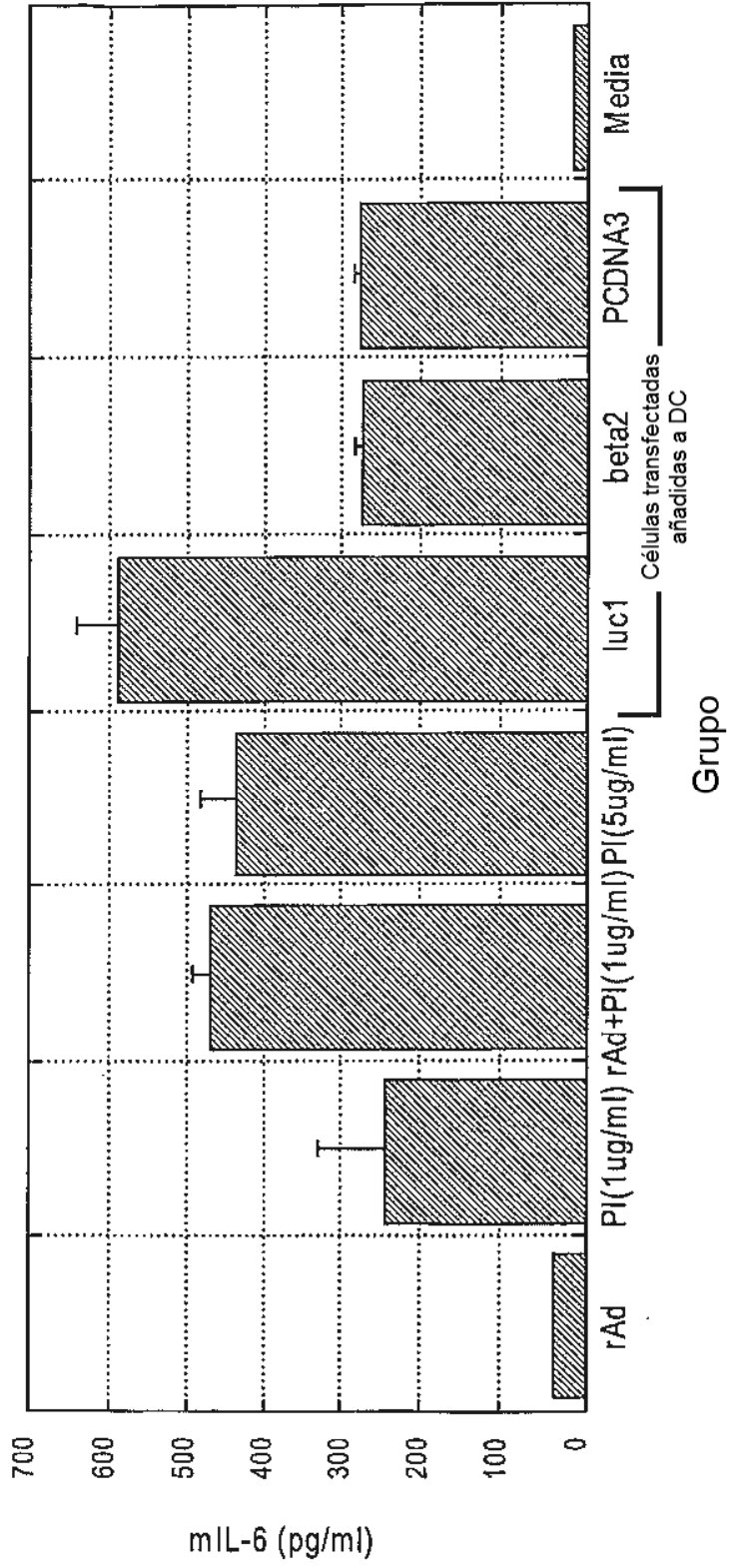


FIG. 4A

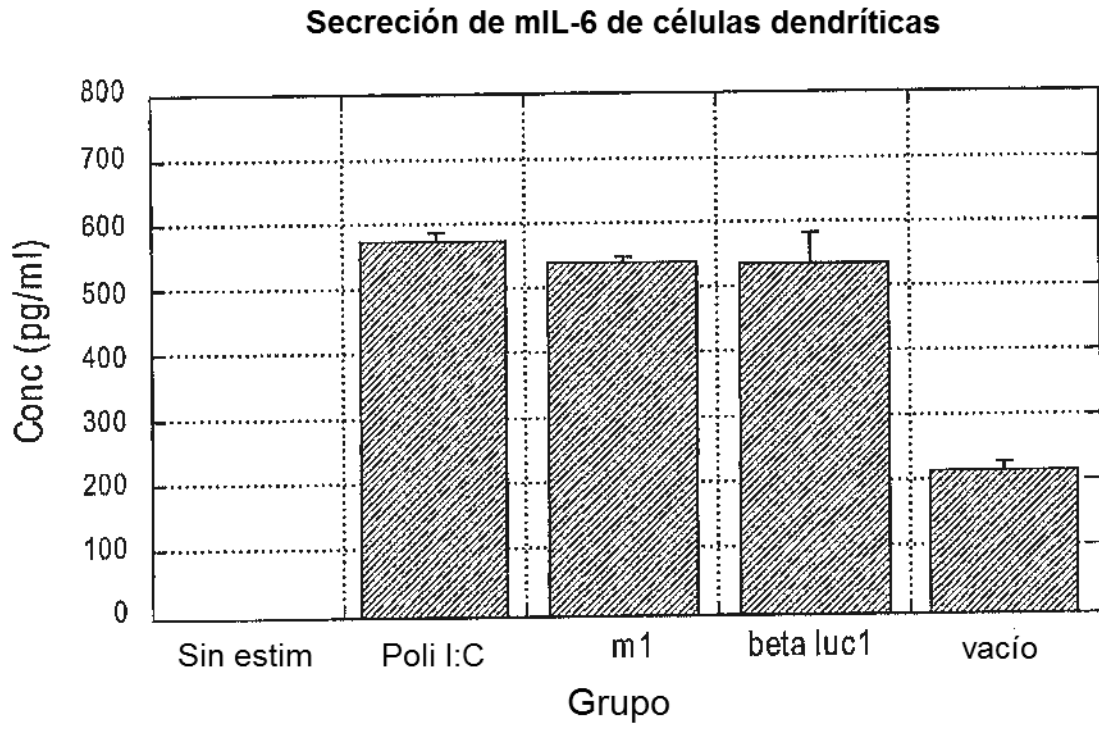
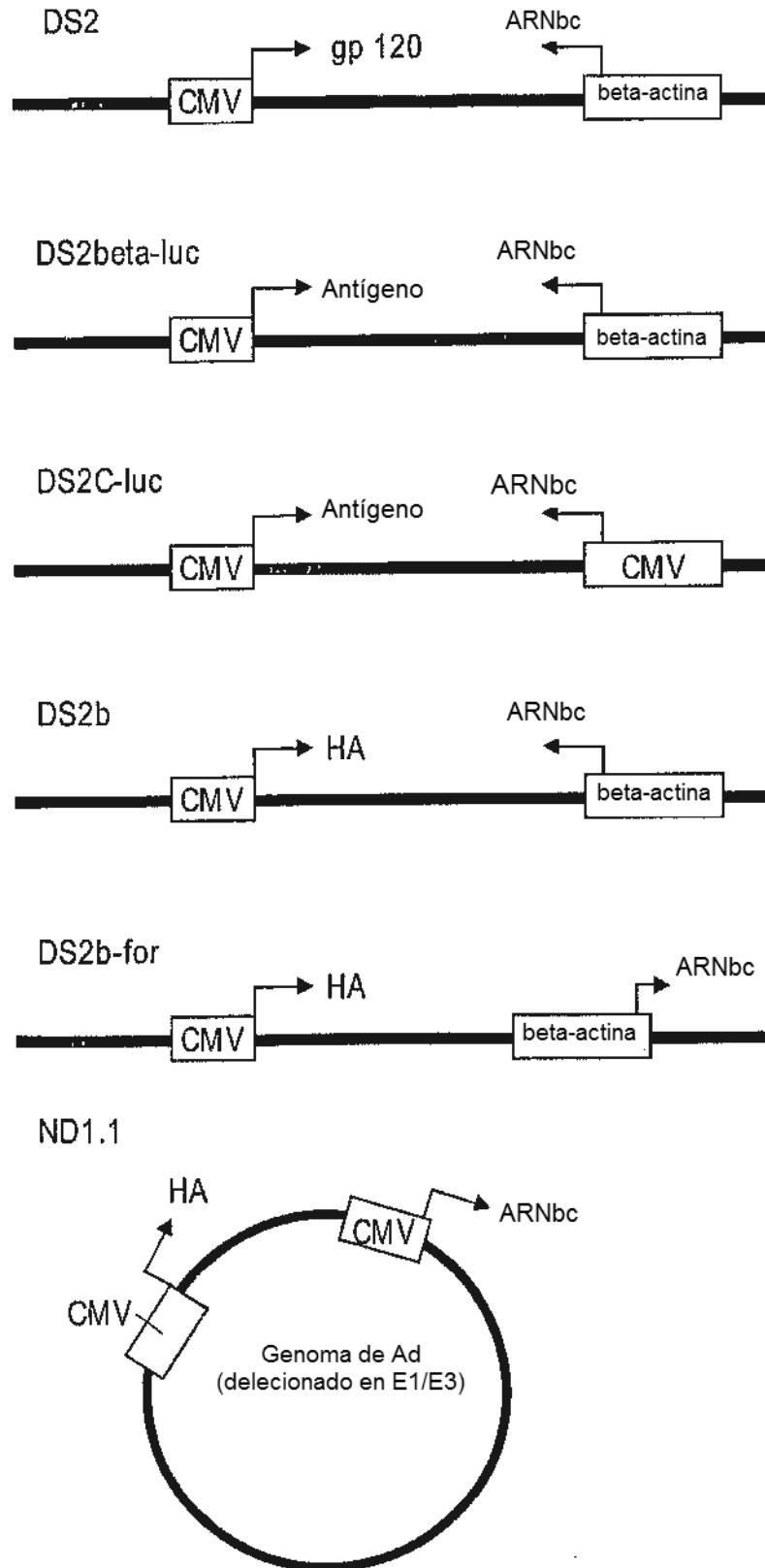


FIG. 4B

FIG. 5

Construcciones de ARNbc expresado



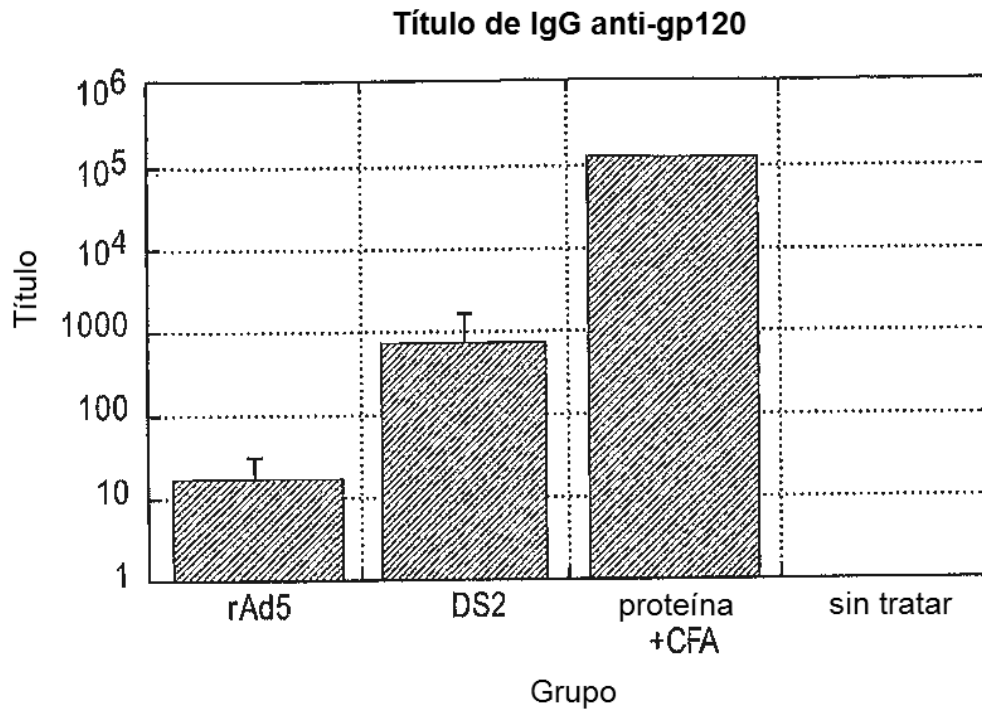


FIG. 6

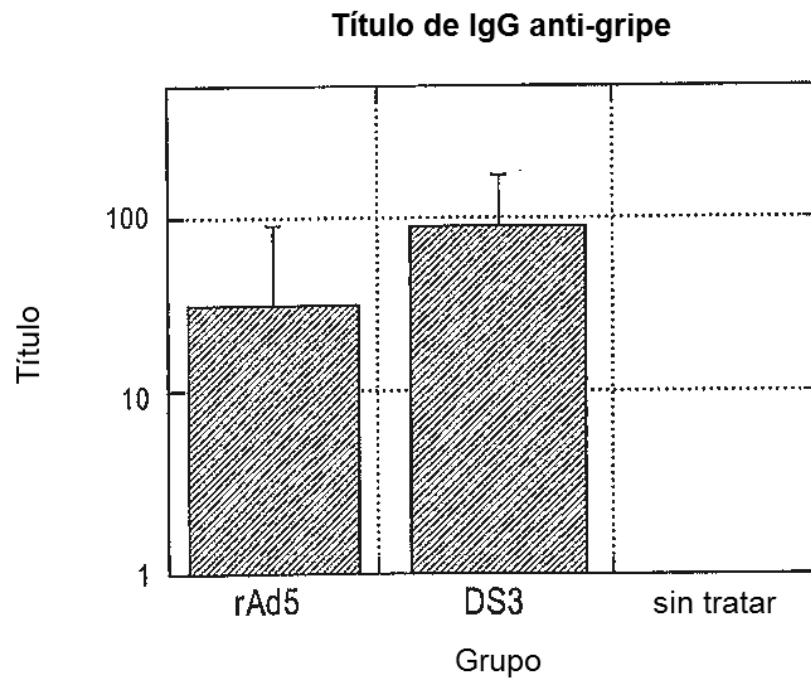


FIG. 7

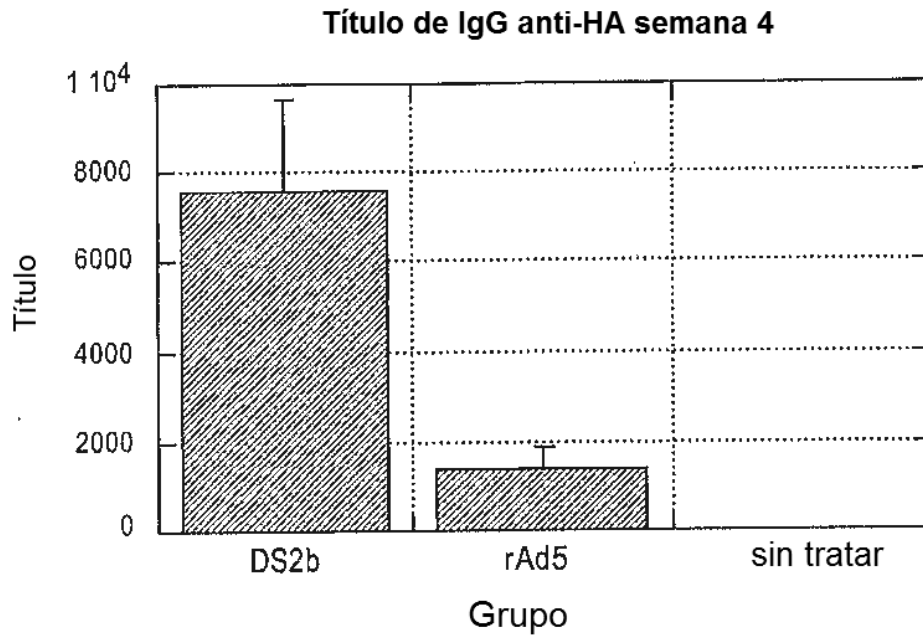


FIG. 8A

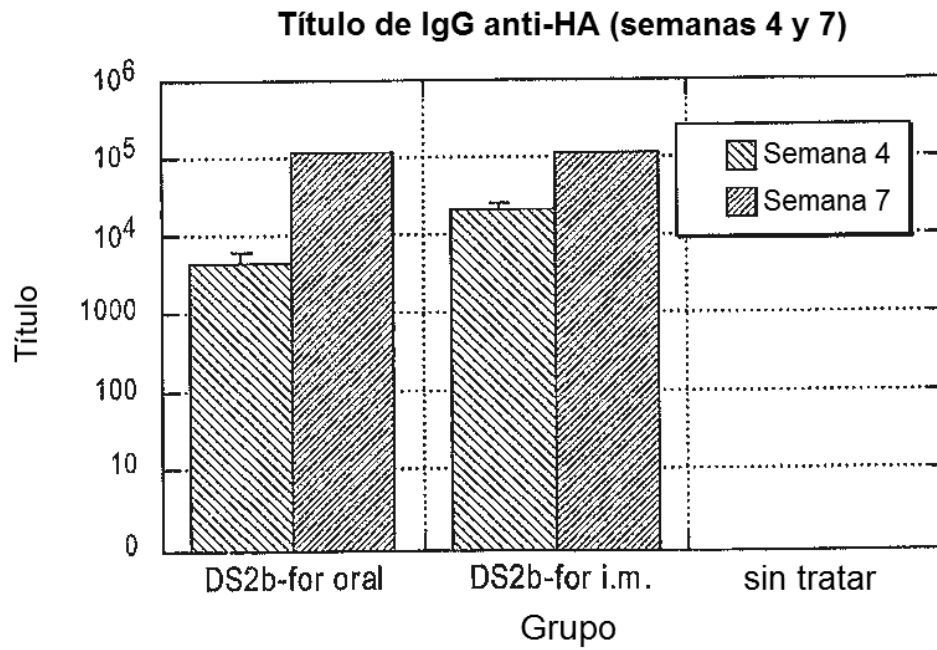


FIG. 8B

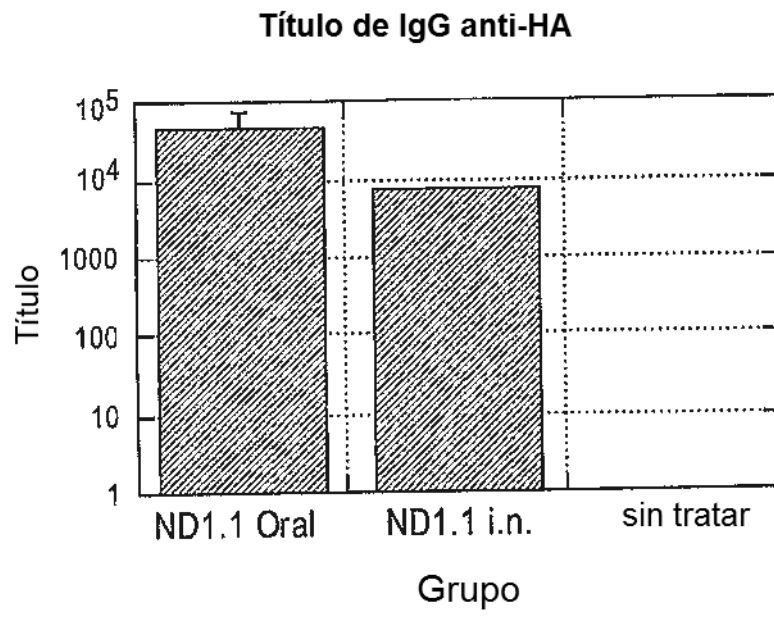


FIG. 8C