

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 928 689**

51 Int. Cl.:

C12N 15/86

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.04.2019** **E 19169121 (1)**

97 Fecha y número de publicación de la concesión europea: **27.07.2022** **EP 3722434**

54 Título: **Sistema de plásmidos**

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
22.11.2022

73 Titular/es:

FREELINE THERAPEUTICS LIMITED (100.0%)
Stevenage Bioscience Catalyst Gunnels Wood
Road
Stevenage, Hertfordshire SG1 2FX, GB

72 Inventor/es:

HÖRER, MARKUS;
SONNTAG, FLORIAN y
KOBER, RENÉE

74 Agente/Representante:

PONS ARIÑO, Ángel

ES 2 928 689 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

Sistema de plásmidos

5 Campo de la invención

La presente invención se refiere a sistemas de dos plásmidos y/o plásmidos auxiliares para producir vectores AAV recombinantes (rAAV). La invención se refiere además a métodos que usan los sistemas de dos plásmidos y/o plásmidos auxiliares de la invención, o a sus usos.

10 Antecedentes de la invención

Los vectores de virus adenoasociados recombinantes ("recombinant adeno-associated virus", rAAV) tienen un potencial considerable para la terapia génica debido a su prometedor perfil de seguridad y su capacidad para transducir muchos tejidos *in vivo*. Sin embargo, la producción sigue siendo bastante difícil y compleja y el aumento de escala de la producción a escala industrial se ha logrado solo en un grado limitado. Una de las razones de esto es que la producción de rAAV depende de una coinfección con un virus auxiliar para propagarse y establecer un ciclo de vida productivo. La infección de células con un virus auxiliar competente para la replicación, por ejemplo, un adenovirus, para la producción de rAAV tiene la desventaja de que las cepas de rAAV resultantes están contaminadas con virus auxiliar, lo cual requiere realizar etapas de eliminación de virus validados en el proceso de purificación posterior.

Por este motivo, el uso de la coinfección por adenovirus suele evitarse proporcionando las funciones auxiliares adenovirales pertinente en un plásmido que se cotransfecta junto con varios otros plásmidos que contienen las funciones Rep y Cap de AAV, y el "genoma del vector" de rAAV, es decir, el ácido nucleico heterólogo que comprende la "carga útil" genética del rAAV, flanqueado por las secuencias de repetición terminal invertida (ITR) que aseguran la encapsidación del ácido nucleico heterólogo dentro de las partículas virales producidas. Por ejemplo, Emmerling *et al.* (2016), en la sección 3.1.2 describe sistemas de cuatro plásmidos en los que las funciones Rep, Cap y auxiliar adenoviral se proporcionan cada una en plásmidos separados, junto con el genoma del vector en un cuarto plásmido. La síntesis de plásmidos es un factor importante del aumento del coste de los productos, y el uso de cuatro plásmidos, todos los cuales deben ingresar a la misma célula para que la producción de rAAV se produzca en esa célula, es una desventaja desde un punto de vista económico y de eficacia.

En la sección 3.1.3 de Emmerling *et al.* (2016) se menciona un sistema de dos plásmidos, en el que un plásmido contiene Rep, Cap y el genoma del vector, proporcionándose las funciones adenovirales en el segundo plásmido. En este sentido, debe señalarse que una de las principales preocupaciones de las agencias reguladoras de medicamentos con respecto a la seguridad de los productos basados en virus, ya sea para virus oncolíticos o vectores de transferencia de genes, es la generación de retromutantes similares al tipo salvaje mediante la recombinación de la información genética contenida en los materiales de partida, por ejemplo, en este caso, tras la transfección de las células de producción con plásmidos que portan colectivamente esta información. Los acontecimientos de recombinación entre los plásmidos pueden dar como resultado la generación de los denominados virus competentes para la replicación ("replication competent", rc) (en el caso de AAV, de partículas de AAV competentes para la replicación (rcAAV)). Cuando Rep y Cap están presentes en el mismo plásmido, existe el riesgo de que se produzcan niveles inaceptables de rcAAV por recombinación intramolecular o intermolecular (dependiendo de qué plásmido porta el ITR-ácido nucleico heterólogo-ITR).

Tal disposición de plásmido también es económicamente subóptima cuando se desea cambiar a un genoma de vector diferente (es decir, un ácido nucleico heterólogo de interés; módulo de transgén), o para cambiar a un serotipo de cápside diferente (es decir, un gen Cap diferente) que podría ser deseable en el caso de que se busque un tropismo de tejido diferente. En cualquier caso, será necesario sintetizar un nuevo plásmido (Rep-Cap-genoma del vector), siendo el coste de cada síntesis en parte una función de la longitud del plásmido, a lo que contribuye el gen Rep (invariable).

A modo de antecedentes adicionales: (i) el documento US 2005/080027 divulga la producción optimizada de vectores virales derivados de parovirus en células de producción y encapsidación mediante infección por hsv o tratamiento con inhibidores de metilación del ADN; (ii) el documento US 2003/103939 divulga virus adenoasociados seudotipados y sus usos; y (iii) el documento CN 108048483 describe un sistema de vector HAdV-5 de adenovirus recombinante de tipo replicación y su aplicación.

Con el fin de satisfacer la demanda actual de material de vector rAAV para ensayos clínicos y suministros para el mercado, los siguientes objetivos aún no se han logrado: (a) la mejora del perfil de seguridad y calidad de los vectores rAAV para cumplir las exigencias reglamentarias sobre la calidad de los vectores; (b) una economía mejorada de la fabricación de rAAV en términos del coste de los materiales de partida para una campaña de fabricación determinada, así como el coste del cambio entre campañas; (c) unos altos rendimientos de producción de rAAV; y (d) el control sobre la producción de partículas de rAAV producidas que contienen ("llenas") o que carecen ("vacías") de un genoma de vector recombinante completo.

Sumario de la invención

La presente invención se refiere a un sistema de dos plásmidos para la fabricación de rAAV, en el que al menos algunas de las funciones del gen auxiliar de adenovirus se combinan con Rep de AAV en un plásmido, y Cap de AAV se combina con el genoma del vector rAAV en un segundo plásmido. El sistema de dos plásmidos "Rep-Cap transdivididos" es mejor que los sistemas conocidos, puesto que logra una combinación de economía mejorada que surge del formato más sencillo y flexible, con atributos de seguridad y calidad mejorados, además de otras ventajas sorprendentes como se describe en este documento.

En consecuencia, la invención incluye los siguientes aspectos:

[1] Un plásmido auxiliar que comprende al menos un gen rep de virus adenoasociado (AAV) que codifica al menos una proteína Rep de AAV funcional y al menos un gen de virus auxiliar, y que no comprende un gen cap que codifica un conjunto funcional de proteínas Cap.

[2] Un sistema de dos plásmidos que comprende el plásmido auxiliar de [1] y un vector plasmídico.

[3] El sistema de dos plásmidos de [2], en el que:

(i) el vector plasmídico comprende:

(a) un gen cap de AAV que codifica al menos una proteína Cap de AAV funcional; o

(b) al menos un promotor del gen cap de AAV, un sitio de clonación unido operativamente al promotor del gen cap de AAV y un módulo de expresión flanqueado en al menos un lado por una repetición terminal invertida ("inverted terminal repeat", ITR);

en el que el vector plasmídico no comprende un gen rep que codifica una proteína Rep funcional y el módulo de expresión comprende un transgén unido operativamente al menos a un elemento de control regulador; y/o

(ii) la proporción de plásmido auxiliar a vector plasmídico está entre 1:1 y 1:4.

[4] El sistema de dos plásmidos o plásmido auxiliar de uno cualquiera de los aspectos anteriores, en el que dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 52 funcional, al menos un gen que codifica una proteína Rep 40 funcional y un gen que codifica una proteína Rep 68 funcional.

[5] El sistema de dos plásmidos o plásmido auxiliar de uno cualquiera de los aspectos anteriores, en el que:

(i) dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 52 funcional, y el gen que codifica una proteína Rep 52 funcional comprende una secuencia de ácido nucleico que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 1000 nucleótidos de longitud de los nucleótidos 993-2186 de SEQ ID NO:1, o con un tramo de nucleótidos correspondiente en un serotipo diferente de AAV; y/o

(ii)

(a) dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 40 funcional, y el gen que codifica una proteína Rep 40 funcional comprende una secuencia de ácido nucleico que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 800 nucleótidos de longitud de un tramo de nucleótidos correspondiente a los nucleótidos 993-2252 menos los nucleótidos 1907-2227 de SEQ ID NO:1, o con tramos de nucleótidos correspondientes en un serotipo diferente de AAV; y/o

(b) dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 40 funcional, y el gen que codifica una proteína Rep 40 funcional comprende una secuencia de ácido nucleico que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 1100 nucleótidos de longitud de los nucleótidos 993-2252 de SEQ ID NO:1, o con un tramo de nucleótidos correspondiente en un serotipo diferente de AAV; y/o

(iii) dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 68 funcional, y el gen que codifica una proteína Rep 68 funcional comprende una secuencia de ácido nucleico que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 1500 nucleótidos de longitud de un tramo de nucleótidos correspondiente a los nucleótidos 321-2252 menos los nucleótidos 1907-2227 de SEQ ID NO:1, o con tramos de nucleótidos correspondientes en un serotipo diferente de AAV.

[6] El sistema de dos plásmidos o plásmido auxiliar de uno cualquiera de los aspectos anteriores, en el que:

(i) el plásmido auxiliar no comprende un gen que codifica una proteína Rep 78 funcional; y/o

(ii) el plásmido auxiliar no comprende una secuencia contigua de al menos 1800 nucleótidos

correspondientes a un tramo de nucleótidos contiguo de longitud equivalente comprendido dentro de los nucleótidos 321-2186 de SEQ ID NO:1 o dentro de un tramo de nucleótidos correspondiente en un serotipo diferente de AAV; y/o

(iii) dicho al menos un gen rep de AAV no comprende un promotor p40 interno funcional; y/o

5 (iv) el plásmido auxiliar no comprende un tramo contiguo de la secuencia del gen cap exclusivamente de más de 60 nucleótidos.

[7] El sistema de dos plásmidos o plásmido auxiliar de uno cualquiera de los aspectos anteriores, en el que:

10 (i) el plásmido auxiliar comprende al menos un gen de virus auxiliar y dicho al menos un gen de virus auxiliar comprende un ácido nucleico asociado a virus ("viral associated", VA), un gen E2A y un gen E4; y/o

(ii) el plásmido auxiliar comprende al menos un gen de virus auxiliar y dicho al menos un gen de virus auxiliar comprende un ácido nucleico VA, un gen E2A y un gen E4, y el gen E4 no está ubicado entre el ácido nucleico VA y el gen E2A; y/o

15 (iii) el plásmido auxiliar comprende al menos un gen de virus auxiliar y el plásmido auxiliar tiene una longitud de entre 12000 pb y 15000 pb; y/o

(iv) el plásmido auxiliar no comprende una secuencia contigua de al menos 3000 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 194-3620 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus; y/o

20 (v) el plásmido auxiliar no comprende una secuencia contigua de al menos 60 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 4032-4100 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus; y/o

(vi) el plásmido auxiliar no comprende una secuencia contigua de al menos 350 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 4051-4413 de SEQ ID NO:1, o un tramo de nucleótidos correspondiente en un serotipo diferente de AAV.

25

[8] El sistema de dos plásmidos o plásmido auxiliar de uno cualquiera de los aspectos anteriores, en el que:

30 (i) el plásmido auxiliar no comprende una secuencia contigua de al menos 600 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 2301-2947 de SEQ ID NO:1, o un tramo de nucleótidos correspondiente en un serotipo diferente de AAV; y/o

(ii) el plásmido auxiliar comprende al menos un gen de virus auxiliar y dicho al menos un gen de virus auxiliar comprende un ácido nucleico VA, un gen E2A y un gen E4, y el ácido nucleico VA, el gen E2A y el gen E4 están comprendidos dentro de una porción contigua de menos de 10000 nucleótidos; y/o

35 (iii) el plásmido auxiliar no comprende una secuencia contigua de al menos 22000 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 10619-32756 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus; y/o

(iv) el plásmido auxiliar y/o el vector plasmídico no comprende un sitio de unión a Rep artificial.

40 [9] El sistema de dos plásmidos o plásmido auxiliar de uno cualquiera de los aspectos anteriores, en el que el plásmido auxiliar comprende al menos un gen del virus auxiliar y dicho al menos un gen del virus auxiliar está comprendido en un tramo contiguo del plásmido que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 8000 nucleótidos de longitud de SEQ ID NO:4.

45 [10] El sistema de dos plásmidos de uno cualquiera de [2]-[9], en el que:

(i) el vector plasmídico comprende al menos un promotor del gen cap de AAV, que comprende un promotor p40 de AAV, un promotor p5 y un promotor p19; y/o

(ii) el vector plasmídico no comprende codones de iniciación de la traducción prescindibles; y/o

50 (iii) el vector plasmídico comprende una región promotora que comprende uno o más promotores, y en la región promotora:

(a) los nucleótidos correspondientes a los nucleótidos 321-323 de SEQ ID NO:1 están ausentes;

(b) los nucleótidos correspondientes a los nucleótidos 766-768 de SEQ ID NO:1 no son ATG;

55 (c) los nucleótidos correspondientes a los nucleótidos 955-957 de SEQ ID NO:1 están ausentes;

(d) los nucleótidos correspondientes a los nucleótidos 993-995 de SEQ ID NO:1 están ausentes; y

(e) los nucleótidos correspondientes a los nucleótidos 1014-1016 de SEQ ID NO:1 están ausentes; y/o

(iv) el vector plasmídico comprende un esqueleto de menos de 3000 nucleótidos de longitud.

60 [11] El sistema de dos plásmidos de [10(iii)(b)], en el que los nucleótidos correspondientes a los nucleótidos 766-768 de SEQ ID NO:1 son ATT.

[12] Un método para producir una preparación de AAV recombinante que comprende:

65 (a) obtener el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico como se define en uno

cualquiera de [1]-[11];

(b) transfectar una célula huésped con el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico como se define en uno cualquiera de [1]-[11]; y

(c) cultivar la célula huésped en condiciones adecuadas para la producción de AAV recombinante.

[13] El método de [12]:

(i) que comprende además una etapa de recolección del AAV recombinante para proporcionar una preparación de AAV recombinante; y/o

(ii) en el que la preparación de AAV recombinante comprende un nivel bajo de AAV competente para la replicación (rcAAV); y/o

(iii) en el que el nivel de rcAAV medido es menor que 1 rcAAV en 10^7 AAV recombinantes, menor que 1 rcAAV en 10^9 AAV recombinantes, o menor que 1 rcAAV en 10^{10} AAV recombinantes; y/o

(iv) en el que la preparación de AAV recombinante comprende un nivel bajo de rcAAV y el nivel de rcAAV es menor que el nivel de rcAAV producido utilizando un método equivalente, excepto que el vector plasmídico comprende al menos un gen rep de AAV y al menos un gen cap de AAV; y/o

(v) comprende una etapa de seleccionar una proporción de plásmido auxiliar a vector plasmídico en el que la proporción de plásmido auxiliar a vector plasmídico se selecciona o ajusta a una proporción que permite al usuario obtener la proporción deseada de partículas llenas a totales o el rendimiento alto o deseado del vector AAV recombinante; y/o

(vi) en el que la proporción de plásmido auxiliar a vector plasmídico se selecciona o ajusta a una proporción que logra un rendimiento máximo de AAV recombinante con el rendimiento mínimo de partículas vacías alcanzable con dicho rendimiento máximo de AAV recombinante; y/o

(vii) en el que la proporción deseada de partículas llenas a totales es una proporción de partículas llenas a totales que es al menos el 20 % o al menos el 30 % de la proporción de partículas llenas a totales lograda utilizando un método equivalente con una proporción de plásmido auxiliar al vector plasmídico de 1,8:1; y/o

(viii) que comprende además una etapa de purificación del AAV recombinante.

Descripción de las figuras

La figura 1 proporciona un esquema del genoma de AAV nativo que muestra transcritos de Rep y Cap (VP1-3) solapantes y la posición de p5, promotores p19 y p40. ITR = repetición terminal invertida.

La figura 2 proporciona un esquema del plásmido auxiliar construido como se describe en el ejemplo 1, con un recuadro que muestra los genes rep, incluidos p5, promotores p19 y p40. Las cruces sobre "p40" indican que estos promotores se han vuelto no funcionales. El área rayada en el gen rep 52/40 indica la presencia de la secuencia del intrón que se corta y empalma en el transcrito rep 52, pero se corta fuera del transcrito rep 40. Ori = origen de replicación bacteriano. KanR = gen de resistencia a la kanamicina. Obsérvese que las características respectivas del plásmido no se muestran a escala.

La figura 3 proporciona un esquema del vector plasmídico construido como se describe en el esquema del plásmido auxiliar construido como ejemplo 1, con un recuadro que muestra el gen cap y la región promotora secuencia arriba que incluye los promotores p5, p19 y p40. Las cruces sobre "ATG" y "GTG" indican que estos posibles codones de iniciación de la traducción se han deletado. Ori = origen de replicación bacteriano. KanR = gen de resistencia a la kanamicina. ITR = repetición terminal invertida. Obsérvese que las características respectivas del plásmido no se muestran a escala.

La figura 4 proporciona esquemas que muestran vectores plasmídicos: (A) contiene un gen cap y un sitio de clonación múltiple, flanqueado por ITR, para la clonación en un módulo de expresión; (B) contiene un gen cap y un sitio de clonación múltiple para la clonación en un módulo de expresión flanqueado por ITR; (C) contiene un módulo de expresión, flanqueado por ITR, y un sitio de clonación múltiple, secuencia abajo de una región promotora que contiene los promotores p5, p19 y p40, para la clonación en un gen cap; (D) contiene un sitio de clonación múltiple, flanqueado por ITR, para la clonación en un módulo de expresión y otro sitio de clonación múltiple, secuencia abajo de una región promotora que contiene los promotores p5, p19 y p40, para la clonación en un gen cap; (E) contiene un sitio de clonación múltiple para la clonación en un módulo de expresión flanqueado por ITR y otro sitio de clonación múltiple, secuencia abajo de una región promotora que contiene los promotores p5, p19 y p40, para la clonación en un gen cap.

La figura 5 proporciona los resultados del análisis de rAAV producido y ensayado como se describe en el ejemplo 2. (A) Título de partículas por ml medido mediante ELISA anticápside en rAAV producido usando el sistema de dos plásmidos en proporciones variables de plásmido auxiliar:vector plasmídico a niveles constantes de ADN plasmídico total. "no divididos" = sistema de dos plásmidos con funciones Rep y Cap en el mismo plásmido, utilizado en relación de plásmido módulo de expresión auxiliar de AdV:rep-cap 1,6:1. "Ctrl" = control negativo: plásmido auxiliar transfectado con pUC19 en lugar de vector plasmídico. (B) Título del genoma del vector (vg) por ml medido por qPCR que amplifica una secuencia dentro del promotor del módulo de expresión en las muestras de rAAV de (A). (C) Proporción de vg (de B) a las partículas totales (de A), expresada como "% de llenas", es

decir, el número de vg como porcentaje del número de partículas. Nota: $1,0E+12 = 1,0 \times 10^{12}$. Las barras de error muestran la desviación estándar del análisis por triplicado de las muestras.

La figura 6 muestra la cuantificación de rAAV en diferentes lotes de rAAV. Se analizó el contenido de rAAV en tres lotes diferentes de rAAV que contenían el mismo módulo de transgén (que codifica el factor IX).

La figura 7 muestra la modulación de los rendimientos y la proporción de partículas llenas a totales mediante la modificación de la proporción de plásmidos. Se ensayaron seis proporciones de plásmidos diferentes para el sistema de dos plásmidos en la producción de rAAV. Las proporciones molares de plásmido de A a F son 3:1, 1,8:1, 1:1,5, 1:2, 1:3 y 1:5, respectivamente. Los rendimientos del genoma del virus se determinaron mediante qPCR específica del módulo del transgén. Los rendimientos de la cápside se determinaron mediante ELISA específico de cápside. Las proporciones de partículas llenas a totales se calcularon en función de los resultados de qPCR y ELISA.

La figura 8 muestra los rendimientos del genoma del virus para cuatro transgenes diferentes en el sistema de encapsidación de dos plásmidos. Se usaron cuatro transgenes diferentes para la encapsidación de rAAV. Se realizaron dos experimentos de encapsidación independientes para cada transgén. Los rendimientos se cuantificaron utilizando qPCR específicas de módulo de transgén.

La figura 9 muestra la producción de rAAV aplicando diferentes serotipos de cap o variantes de cap sintéticas. El rAAV se generó aplicando la proporción de plásmido molar idéntica, cinco serotipos de cap diferentes/variantes de cap sintéticas y el mismo módulo transgénico. Los rendimientos del genoma del virus se determinaron mediante qPCR específica de módulo de transgén.

Descripción detallada

Definiciones generales

A menos que se definan de otra manera, los términos y las expresiones científicos y técnicos utilizados en el presente documento tienen el mismo significado que entienden normalmente los expertos en la materia a la que pertenece la presente invención.

En general, la expresión "que incluye" se utiliza para significar "que incluye, entre otros/otras". Por ejemplo, la expresión "un plásmido auxiliar que comprende al menos un gen rep" debe interpretarse en el sentido de que el plásmido auxiliar tiene al menos un gen rep, pero el plásmido auxiliar puede comprender componentes adicionales, tales como genes adicionales.

En algunas realizaciones de la invención, la expresión "que comprende" reemplaza a la expresión "que consiste en" o la expresión "que consiste fundamentalmente en". La expresión "que consiste en" pretende ser limitante. Por ejemplo, la expresión "un plásmido auxiliar que consiste en un gen rep" debe entenderse que significa que el plásmido auxiliar tiene un gen rep y ningún otro material genético. De forma análoga, la expresión "un plásmido auxiliar que consiste fundamentalmente en un gen rep" debe entenderse que significa que el plásmido auxiliar tiene un gen rep y no comprende componentes adicionales que afecten materialmente a la función del plásmido auxiliar. Por ejemplo, un plásmido auxiliar que consiste fundamentalmente en un gen rep no contendrá ningún otro gen, pero podría contener otro material genético, como espaciadores.

Los términos "proteína" y "polipéptido" se usan indistintamente en este documento y pretenden referirse a una cadena polimérica de aminoácidos de cualquier longitud.

Para los fines de la presente invención, para determinar el porcentaje de identidad de dos secuencias (tales como dos secuencias polinucleotídicas o dos secuencias polipeptídicas), las secuencias se alinean con fines de una comparación óptima (por ejemplo, pueden introducirse huecos en una primera secuencia para un alineamiento óptimo con una segunda secuencia). Luego se comparan los nucleótidos o los aminoácidos en cada posición. Cuando una posición en la primera secuencia está ocupada por el mismo aminoácido o nucleótido que la posición correspondiente en la segunda secuencia, entonces los aminoácidos o los nucleótidos son idénticos en esa posición. El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias (es decir, porcentaje de identidad = número de posiciones idénticas/número total de posiciones en la secuencia de referencia x 100).

Normalmente, la comparación de secuencias se lleva a cabo a lo largo de la secuencia de referencia. Por ejemplo, si el usuario desea determinar si una secuencia concreta ("de ensayo") es 95 % idéntica a SEQ ID NO:1, SEQ ID NO:1 sería la secuencia de referencia. Para evaluar si una secuencia es al menos un 80 % idéntica a la SEQ ID NO:1 (un ejemplo de una secuencia de referencia), los expertos llevarían a cabo un alineamiento a lo largo de la SEQ ID NO:1 e identificarían cuántas posiciones en la secuencia de ensayo son idénticas a las de SEQ ID NO:1. Si al menos el 80 % de las posiciones son idénticas, la secuencia de ensayo es al menos el 80 % idéntica a la SEQ ID NO:1. Si la secuencia es más corta que la SEQ ID NO:1, los huecos o posiciones que faltan deben considerarse posiciones no

idénticas.

Los expertos en la materia conocen diferentes programas informáticos que están disponibles para determinar la homología o la identidad entre dos secuencias. Por ejemplo, puede realizarse la comparación de secuencias y la determinación del porcentaje de identidad entre dos secuencias mediante un algoritmo matemático. En una realización, el porcentaje de identidad entre dos secuencias de aminoácidos o de ácidos nucleicos se determina utilizando el algoritmo de Needleman y Wunsch (1970) que se ha incorporado al programa GAP en el paquete de software Accelrys GCG (disponible en <http://www.accelrys.com/products/gcg/>), usando una matriz Blosum 62 o una matriz PAM250 y una ponderación por hueco de 16, 14, 12, 10, 8, 6 o 4 y una ponderación por longitud de 1, 2, 3, 4, 5 o 6.

En el presente documento, el término "plásmido" pretende indicar una molécula de ácido nucleico que puede replicarse independientemente de un cromosoma celular. El término "plásmido" pretende abarcar moléculas de ácido nucleico circulares y moléculas de ácido nucleico lineales. Además, el término "plásmido" pretende abarcar plásmidos bacterianos, pero también cósmidos, minicírculos (Nehlsen, K., Broll S., Bode, J. (2006), Gene Ther. Mol. Biol., 10:233-244; Kay, M.A., He, C.-Y., Chen, Z.-N. (2010), Nature Biotechnology, 28:1287-1289) y minicuerdas (Nafissi N., Alqawlaq S., Lee E.A., Foldvari M., Spagnuolo P.A., Slavcev R.A. (2014), Mol. Ther. Nucleic Acids, 3:e165). Opcionalmente, el plásmido es una molécula de ácido nucleico circular. Opcionalmente, el plásmido es una molécula de ácido nucleico de origen bacteriano.

El término "auxiliar" no pretende ser limitante. En consecuencia, un "plásmido auxiliar" es cualquier plásmido que:

(i) comprende al menos un gen rep que codifica al menos una proteína Rep funcional y no comprende un gen cap que codifica una proteína Cap funcional; o

(ii) comprende al menos un gen de virus auxiliar y no comprende un gen cap que codifica una proteína Cap funcional, y dicho al menos un gen de virus auxiliar está comprendido en un tramo contiguo del plásmido que tiene al menos un 95 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad con la longitud completa o con un fragmento de al menos 6000, al menos 7000 o al menos 8000 nucleótidos de longitud de SEQ ID NO:4.

La expresión "vector plasmídico" no pretende ser limitante. En consecuencia, un "vector plasmídico" es cualquier plásmido que:

(i) es adecuado para usar junto con un plásmido auxiliar en un sistema de dos plásmidos de la invención; o

(ii) comprende:

(a) un gen cap que codifica al menos una proteína Cap funcional; o

(b) al menos un promotor del gen cap, un sitio de clonación unido operativamente al promotor del gen cap y un módulo de expresión flanqueado en al menos un lado por una ITR;

en el que el vector plasmídico no comprende un gen rep que codifica una proteína Rep funcional y el módulo de expresión comprende un transgén unido operativamente al a menos un elemento de control regulador.

La expresión "molécula de ácido nucleico" se refiere a una forma polimérica de nucleótidos de cualquier longitud. Los nucleótidos pueden ser desoxirribonucleótidos, ribonucleótidos o sus análogos. Preferentemente, el plásmido está formado por desoxirribonucleótidos o ribonucleótidos. Incluso más preferentemente, el plásmido está formado por desoxirribonucleótidos, es decir, el plásmido es una molécula de ADN.

La expresión "tipo salvaje" y el término "nativo" son sinónimos y se refieren a genes presentes en el genoma de una cepa/serotipo de AAV o adenovirus, o a proteínas codificadas por genes presentes en el genoma de una cepa/serotipo de AAV o adenovirus.

Ensayo de producción de AAV

En un ensayo de producción de AAV, el usuario puede determinar si un determinado plásmido o sistema de dos plásmidos de "ensayo" son eficaces para producir rAAV a un nivel similar a un plásmido o sistema de dos plásmidos de "referencia" como sigue.

El usuario proporciona un sistema de dos plásmidos de "referencia" que comprende un "plásmido auxiliar de referencia" y un "vector plasmídico de referencia".

Por ejemplo, el usuario proporciona un plásmido auxiliar de referencia que comprende genes auxiliares de adenovirus 5 de tipo salvaje que codifican E2A, E4 y VA ARN I y II, es decir, los genes auxiliares de adenovirus comprendidos en SEQ ID NO:2. Los detalles de las posiciones de los ácidos nucleicos en SEQ ID NO:2 que codifican estos genes se exponen con más detalle a continuación bajo el encabezado "Al menos un gen auxiliar de virus". El plásmido auxiliar también comprende un gen rep de tipo salvaje que codifica Rep 40, Rep 52, Rep 68 y Rep 78 y los promotores de rep p5, p19 y p40, es decir, las secuencias comprendidas dentro de los nucleótidos 200-

2252 de SEQ ID NO:1.

El usuario proporciona un vector plasmídico de referencia que comprende un gen cap de tipo salvaje unido operativamente a un promotor del gen cap de tipo salvaje que comprende p5, p19 y p40, es decir, el gen cap comprendido en SEQ ID NO:1 (nucleótidos 5961-8171 de SEQ ID NO:1). El vector plasmídico comprende además un transgén flanqueado por dos ITR de AAV2, es decir, las ITR comprendidas en los nucleótidos 1-145 y 4535-4679 de SEQ ID NO:1.

A continuación, el usuario proporciona un sistema de dos plásmidos de "ensayo" que comprende un "plásmido auxiliar de ensayo" y un "vector plasmídico de ensayo" que se basan en el sistema de dos plásmidos de referencia, pero tiene un solo cambio relacionado con una característica que los usuarios desean analizar. Por ejemplo, si el usuario desea saber si una determinada proteína rep es funcional, el usuario podría extraer el gen rep del "plásmido auxiliar de referencia" y reemplazarlo con la proteína rep de ensayo para proporcionar un "plásmido auxiliar de ensayo".

Luego, el usuario comparará la capacidad del sistema de dos plásmidos de referencia y el sistema de dos plásmidos de ensayo para permitir la producción de rAAV. Para ello, el usuario puede transfectar un primer conjunto de células huésped adecuadas (como las células HEK293T que expresan el gen E1A/B, de modo que el plásmido auxiliar no necesita comprender un gen E1A/B) con el sistema de dos plásmidos de referencia, y un segundo conjunto de células huésped idénticas con el sistema de dos plásmidos de ensayo e incubar las células huésped durante un período de tiempo adecuado para que se produzca la producción de AAV. El rendimiento del AAV recombinante producido a partir del sistema de dos plásmidos de referencia y el sistema de dos plásmidos de ensayo se puede recolectar y medir mediante qPCR para cuantificar el número de genomas del vector. Por ejemplo, puede usarse qPCR para determinar el número de casos de moléculas de ácido nucleico que comprenden un componente del genoma del vector, como una secuencia promotora, que se producen en las células huésped transfectadas con el sistema de dos plásmidos de ensayo en comparación con las células huésped transfectadas con el sistema de dos plásmidos de referencia. Como alternativa, se puede determinar el rendimiento comparativo de partículas, por ejemplo, mediante un ELISA anticápside.

Un ensayo de producción de AAV adecuado se describe en el ejemplo 2.

Sistema de dos plásmidos

La presente invención proporciona un sistema de dos plásmidos que comprende un plásmido auxiliar y un vector plasmídico, en el que el plásmido auxiliar comprende al menos un gen rep de AAV que codifica al menos una proteína Rep de AAV funcional y al menos un gen de virus auxiliar, y no comprende un gen cap que codifica una proteína Cap funcional.

El sistema de dos plásmidos es útil para producir rAAV. Opcionalmente, el sistema de dos plásmidos es adecuado para su uso en la producción de rAAV. Opcionalmente, el sistema de dos plásmidos es para producir rAAV. Opcionalmente, el sistema de dos plásmidos es para producir rAAV adecuados para su uso en terapia génica. Opcionalmente, el sistema de dos plásmidos es para producir rAAV para su uso en terapia génica.

La expresión "sistema de dos plásmidos" se refiere a un sistema que comprende dos plásmidos y puede usarse sin necesidad de plásmidos adicionales para producir rAAV. Opcionalmente, el sistema de dos plásmidos se puede utilizar para producir rAAV sin necesidad de un virus auxiliar, como adenovirus. Opcionalmente, el sistema de dos plásmidos se puede utilizar para producir rAAV sin necesidad de material genético procedente de una célula huésped, opcionalmente con la excepción de un gen que codifica E1A/B. Sin embargo, el sistema puede comprender componentes adicionales que no sean plásmidos. Opcionalmente, el sistema de dos plásmidos no comprende un virus auxiliar. Opcionalmente, el sistema de dos plásmidos de la invención comprende toda la información genética necesaria para la producción de rAAV. Por ejemplo, el sistema de dos plásmidos de la invención puede comprender al menos un gen rep de AAV, al menos un gen cap de AAV y al menos un gen auxiliar. Opcionalmente, el sistema de dos plásmidos de la invención comprende toda la información genética necesaria requerida para la producción de rAAV adecuados para su uso en terapia génica. Por ejemplo, el sistema de dos plásmidos de la invención puede comprender al menos un gen rep de AAV, al menos un gen cap de AAV, al menos un gen auxiliar y un módulo de expresión que comprende un transgén unido operativamente al menos a un elemento de control regulador. Sin embargo, en realizaciones, el sistema de dos plásmidos de la invención puede carecer de un gen cap funcional (necesario para la producción de rAAV) y/o un módulo de expresión que comprende un transgén unido operativamente al menos a un elemento de control regulador (necesario para la producción de rAAV adecuados para su uso en terapia génica).

Es una ventaja de la presente invención que un gen cap de AAV y/o un transgén en el plásmido del vector puedan cambiarse por otros para tratar diferentes trastornos genéticos. Opcionalmente, por lo tanto, el sistema de dos plásmidos de la invención puede comprender toda la información genética necesaria para la producción de rAAV excepto un gen cap funcional, y en tales realizaciones, el sistema de dos plásmidos de la invención puede comprender un sitio adecuado para la clonación en un gen cap de AAV. Dicho sitio puede comprender un sitio de

clonación adyacente a un promotor del gen cap de AAV. El sitio adecuado para la clonación en un gen cap de AAV estará presente en el vector plasmídico. Opcionalmente, el sistema de dos plásmidos de la invención comprende toda la información genética necesaria para la producción de rAAV adecuados para su uso en terapia génica excepto un gen cap funcional y un módulo de expresión que comprende un transgén y un elemento de control regulador, en el que dicho sistema de dos plásmidos comprende un sitio adecuado para la clonación en un gen cap de AAV (tal como un promotor del gen cap de AAV y un sitio de clonación unido operativamente, es decir, adyacente, al promotor del gen cap de AAV) y un sitio adecuado para la clonación en un módulo de expresión (tal como un sitio de clonación flanqueado por una o más ITR). Opcionalmente, el sistema de dos plásmidos de la invención comprende toda la información genética necesaria para la producción de rAAV adecuados para su uso en terapia génica excepto un módulo de expresión que comprende un transgén y un elemento de control regulador, en el que dicho sistema de dos plásmidos comprende un sitio adecuado para la clonación en un módulo de expresión (tal como un sitio de clonación flanqueado por una o más ITR). En tales casos, el sitio adecuado para la clonación en un gen cap de AAV y el sitio adecuado para la clonación en un módulo de expresión estarán presentes en el vector plasmídico. Opcionalmente, el sistema de dos plásmidos de la invención comprende los siguientes componentes divididos entre el vector plasmídico y el plásmido auxiliar:

- al menos un gen rep de AAV que codifica al menos una proteína Rep de AAV funcional;
- al menos un gen de virus auxiliar;
- un gen cap de AAV que codifica al menos una proteína funcional de la cápside de AAV, o un promotor del gen cap de AAV y un sitio de clonación unido operativamente al promotor del gen cap de AAV;
- al menos una ITR; y
- un módulo de expresión que comprende un transgén unido operativamente al menos a un elemento de control regulador, o un sitio adecuado para la clonación en un módulo de expresión flanqueado en al menos un lado por una ITR (es decir, adyacente a esta).

El vector plasmídico puede comprender:

- (a) un gen cap de AAV que codifica al menos una proteína Cap de AAV funcional; o
- (b) al menos un promotor del gen cap de AAV, un sitio de clonación unido operativamente al promotor del gen cap de AAV, y un módulo de expresión flanqueado en al menos un lado por una ITR;

en el que el vector plasmídico no comprende un gen rep que codifica una proteína Rep funcional y el módulo de expresión comprende un transgén unido operativamente al a menos un elemento de control regulador.

El plásmido auxiliar comprende al menos un gen rep de AAV que codifica al menos una proteína Rep de AAV funcional y al menos un gen de virus auxiliar, y no comprende un gen cap que codifica una proteína Cap funcional.

Opcionalmente, la proporción de plásmido auxiliar a vector plasmídico es de entre 3:1 y 1:10, entre 1,5:1 y 1:9, entre 1,4:1 y 1:8, entre 1,3:1 y 1:7, entre 1,2:1 y 1:6, entre 1,1:1 y 1:5, entre 1:1 y 1:4, o entre 1:1,5 y 1:3. Opcionalmente, el sistema de dos plásmidos comprende un exceso molar de vector plasmídico en comparación con el plásmido auxiliar.

Como se expone con más detalle a continuación, la alteración de la proporción de plásmido auxiliar a vector plasmídico se puede usar para controlar la proporción de partículas llenas a totales (o la proporción de cápsides/partículas que están llenas) producidas durante la producción de rAAV, o para aumentar el rendimiento de rAAV producido durante la producción de rAAV.

Mediante la expresión "proporción de plásmido auxiliar a vector plasmídico" se entiende una proporción molar, es decir, la proporción entre el número de moles de plásmido auxiliar presente y el número de moles de vector plasmídico presente. Puede proporcionarse una proporción dada de plásmido auxiliar a vector plasmídico simplemente mezclando el plásmido auxiliar y el vector plasmídico en una proporción molar apropiada.

Plásmido auxiliar

La presente invención proporciona un plásmido auxiliar que comprende al menos un gen rep de AAV que codifica al menos una proteína Rep de AAV funcional y al menos un gen de virus auxiliar, y que no comprende un gen cap que codifica una proteína Cap funcional.

La presente invención se refiere a un sistema mejorado de dos plásmidos que es adecuado para producir rAAV. Un plásmido auxiliar que comprende al menos un gen rep de AAV que codifica al menos una proteína Rep de AAV funcional y no comprende un gen cap que codifica una proteína Cap funcional puede ser ventajoso para usar en un sistema de dos plásmidos mejorado de este tipo. Tanto los genes rep como los genes cap son necesarios para producir rAAV. Sin embargo, los plásmidos y los sistemas de dos plásmidos de la invención están dispuestos de tal manera que los genes rep y cap no están ambos presentes en un único plásmido (configuración "transdivisión"). Es difícil asegurarse de que el plásmido auxiliar no comprenda tanto un gen rep como un gen cap, ya que los genes nativos cap y rep de AAV se solapan y, por lo tanto, es difícil separarlos, mientras se mantiene suficiente material

genético para permitir la producción de la proteína completa. Sin embargo, es ventajoso dividir los genes rep y los genes cap entre un plásmido auxiliar y un vector plasmídico, ya que aumenta el número de acontecimientos de recombinación necesarios para formar rcAAV.

- 5 El AAV solo puede propagarse en presencia de un virus auxiliar que codifica proteínas que ayudan en la propagación de AAV. Sin embargo, el cultivo de AAV en presencia de un virus auxiliar no es ventajoso, ya que los virus auxiliares pueden ser líticos para las células, incluidas las células huésped utilizadas para cultivar AAV. Además, si se utilizan virus auxiliares en la producción de productos de rAAV, como rAAV para su uso en terapia génica, el virus auxiliar puede contaminar el producto. Como alternativa a la coinfección con virus auxiliares, como el
- 10 adenovirus, los genes necesarios del virus auxiliar pueden proporcionarse en un plásmido transfectado. Para las células huésped que expresan los genes adenovirales E1A/B (como las células HEK293T), los genes auxiliares adenovirales necesarios restantes codifican E4, E2A y VA ARN I y II. Si bien estos genes están distribuidos a lo largo de un largo tramo del genoma adenoviral, los presentes inventores han determinado que pueden eliminarse grandes tramos de nucleótidos no codificantes que separan los genes genómicos sin afectar a la expresión de los genes. Por lo tanto, los inventores han diseñado una región mínima de gen auxiliar que es SEQ ID NO:4. El uso de una región mínima de gen auxiliar de este tipo en un plásmido para la producción de rAAV es ventajoso, ya que permite al usuario utilizar un plásmido más pequeño que es menos costoso de producir y más fácil de transfectar a la célula.

20 *Al menos un gen rep de AAV*

El plásmido auxiliar comprende al menos un gen rep de AAV que codifica al menos una proteína Rep de AAV funcional. El AAV comprende una región del gen rep de AAV que codifica cuatro proteínas Rep de AAV (Rep 78, Rep 68, Rep 52 y Rep 40). La región del gen está bajo el control de los promotores p5 y p19. Cuando se utiliza el

25 promotor p5, se transcribe un gen que codifica Rep 78 y Rep 68. Rep 78 y Rep 68 son dos variantes de corte y empalme alternativas (Rep 78 comprende un intrón que se elimina en Rep 68). De forma análoga, cuando se usa el promotor p19, se transcribe un gen que codifica Rep 52 y Rep 40. Rep 52 y Rep 40 son variantes de corte y empalme alternativas (Rep 52 comprende un intrón que se elimina en Rep 40).

30 Se sabe que las cuatro proteínas Rep de AAV están implicadas en la replicación y la encapsidación del genoma viral y son, por lo tanto, útiles en la producción de rAAV.

No es necesario que estén presentes las cuatro proteínas Rep de AAV. Opcionalmente, sin embargo, dicho al menos un gen rep de AAV codifica una proteína Rep grande (Rep 78 o Rep 68) y una proteína Rep de AAV pequeña (Rep 52 o Rep 40). Rep 78 puede ser tóxica para las células, y Rep 78 no necesita estar presente para que tenga lugar la replicación de AAV. En realizaciones en las que dicho al menos un gen rep de AAV no codifica Rep 78, dicho al menos un gen rep de AAV preferentemente codifica Rep 68. En consecuencia, el plásmido auxiliar puede comprender al menos un gen rep de AAV que codifica:

- 40 (a) una proteína Rep 52 funcional;
(b) una proteína Rep 40 funcional; y/o
(c) una proteína Rep 68 funcional.

Una proteína Rep de AAV "funcional" es una proteína que permite la producción de partículas de AAV. Concretamente, se cree que Rep 78 o Rep 68 (las proteínas Rep de AAV grandes) están implicadas en la replicación del genoma de AAV, y se cree que Rep 52 y Rep 40 (las proteínas Rep de AAV pequeñas) están implicadas en la encapsidación del genoma de AAV en una cápside. Está dentro de las capacidades de los expertos determinar si una proteína Rep de AAV dada es funcional. Los expertos simplemente necesitan determinar si la proteína Rep de AAV facilita la producción de AAV utilizando un ensayo de producción de AAV como se describió anteriormente. En este caso, el sistema de dos plásmidos de ensayo comprenderá un plásmido auxiliar que comprende la proteína Rep de AAV cuya "funcionalidad" debe determinarse y, por lo demás, el sistema de dos plásmidos de ensayo utilizado será idéntico al sistema de dos plásmidos de referencia.

En una realización, dicho al menos un gen rep de AAV del plásmido auxiliar codifica una proteína Rep de AAV "funcional". Si la proteína Rep de AAV facilita la producción de rAAV en un nivel de al menos el 25 %, al menos el 40 %, al menos el 50 %, al menos el 70 %, al menos el 80 %, al menos el 90 % o al menos el 95 % del nivel facilitado por la proteína Rep AAV de tipo salvaje, es decir, si el rendimiento de los genomas del vector rAAV producidos es de al menos un 25 %, al menos un 40 %, al menos un 50 %, al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % del rendimiento de los genomas del vector rAAV producidos utilizando el

60 sistema de dos plásmidos de referencia. Preferentemente, dicho al menos un gen rep de AAV del plásmido auxiliar es funcional si facilita la producción de rAAV a un nivel de al menos el 80 % del nivel facilitado por la proteína Rep de AAV de tipo salvaje.

En general, una proteína Rep de AAV solo podrá facilitar la producción de rAAV si es compatible con las ITR que rodean el genoma del AAV que se va a encapsidar. Es posible que algunas proteínas Rep de AAV solo puedan encapsidar material genómico (como un módulo de expresión) cuando está flanqueado por una o más ITR del

mismo serotipo que la proteína Rep de AAV. Otras proteínas Rep de AAV son compatibles entre sí, lo que significa que pueden encapsidar material genómico que está flanqueado por una o más ITR de un serotipo diferente. Por ejemplo, en el caso de un sistema de dos plásmidos, se prefiere que la proteína Rep de AAV sea capaz de facilitar la replicación y la encapsidación de un módulo de expresión comprendido dentro del vector plasmídico, y que tal proteína Rep de AAV sea compatible con al menos una ITR que flanquea el módulo de expresión (es decir, que sea capaz de replicarse y encapsidar el módulo de expresión flanqueado al menos en un lado por una ITR).

Opcionalmente, dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 52 funcional, al menos un gen que codifica una proteína Rep 40 funcional y un gen que codifica una proteína Rep 68 funcional.

El plásmido auxiliar puede comprender dos genes que codifican una proteína Rep 40 funcional. En una realización, dicho al menos un gen rep de AAV comprende dos genes que codifican una proteína Rep 40 funcional. Por ejemplo, el plásmido auxiliar puede comprender dos genes rep de AAV que están separados en el plásmido. El primero de los dos genes rep de AAV separados podría codificar Rep 68 (por ejemplo, usando el promotor p5 o un promotor diferente situado cerca de la posición normal del promotor p5 en el gen rep de AAV) y Rep 40 (por ejemplo, usando el promotor p19 o un promotor diferente situado cerca de la posición normal del promotor p19). El segundo de los dos genes rep de AAV separados podría codificar Rep 52 y Rep 40. Rep 52 y Rep 40 son variantes de corte y empalme alternativas.

Si el plásmido auxiliar comprende dos genes que codifican una proteína Rep 40, uno de los dos genes que codifica una proteína Rep 40 funcional puede comprender un intrón. En una realización, ambos genes que codifican una proteína Rep 40 funcional comprenden un intrón. Sin embargo, en una realización preferida, solo uno de los genes que codifica una proteína Rep 40 funcional comprende un intrón. Por ejemplo, si el usuario desea evitar dicho al menos un gen rep de AAV que codifica Rep 78, el gen rep de AAV puede dividirse mediante duplicación parcial en dos genes. Un gen podría comprender nucleótidos correspondientes al gen rep de AAV nativo de longitud completa con la secuencia correspondiente al intrón eliminado. Tal gen codificaría Rep 68 y Rep 40, pero no codificaría ni Rep 78 ni Rep 52, ya que una porción de cada una de las proteínas Rep 78 y Rep 52 es codificada por la secuencia que actúa como un intrón en el contexto de Rep 40. El segundo gen podría comprender nucleótidos correspondientes a la región del gen rep de AAV nativo secuencia abajo del promotor p19, que codificaría Rep 52 (intrón cortado y empalmado) y Rep 40 (intrón eliminado).

SEQ ID NO:1 proporciona la secuencia del genoma de AAV2 de tipo salvaje, y los nucleótidos 321-2252 de SEQ ID NO:1 codifican las cuatro proteínas Rep de AAV. El gen rep de AAV de longitud completa (nucleótidos 321-2252) codifica las cuatro proteínas Rep de AAV (Rep 78 y Rep 68 a partir del promotor p5, y Rep 52 y Rep 40 a partir del promotor p19). Un tramo más corto del gen rep de AAV secuencia abajo del promotor p19 (nucleótidos 993-2252) codifica Rep 52 y Rep 40 solamente (es decir, este tramo del gen rep de AAV se extiende desde el final del promotor p19 hasta el final del gen). Los nucleótidos 1907-2227 de SEQ ID NO:1 corresponden a un intrón. Rep 78 y Rep 52 comprenden aminoácidos codificados por el intrón, pero Rep 68 y Rep 40 son variantes de corte y empalme alternativas que no comprenden aminoácidos codificados por el intrón. La relación entre el gen rep de AAV y las cuatro proteínas Rep de AAV se muestra en la figura 1.

Opcionalmente, el sistema de dos plásmidos o plásmido auxiliar comprende un gen que codifica una proteína Rep 52 funcional, y el gen que codifica una proteína Rep 52 funcional comprende una secuencia de ácido nucleico que tiene al menos un 95 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad con la longitud completa o con un fragmento de al menos 800, al menos 900, al menos 1000 o al menos 1100 nucleótidos de longitud de los nucleótidos 993-2186 de SEQ ID NO:1, o con un tramo de nucleótidos correspondiente en un serotipo diferente de AAV.

Está dentro de las capacidades de los expertos en la materia determinar si un tramo concreto (de ensayo) de nucleótidos es un "tramo de nucleótidos correspondiente en un serotipo diferente de AAV". Todo lo que se requiere es que los expertos en la materia alineen el tramo de nucleótidos de ensayo con el genoma del serotipo de referencia (es decir, SEQ ID NO:1). Si el tramo de nucleótidos de ensayo tiene más del 90 % de identidad con un tramo de nucleótidos contiguo de la misma longitud en SEQ ID NO:1, el tramo contiguo es un tramo de nucleótidos correspondiente en un serotipo diferente de AAV. Lo mismo se aplica en el caso de secuencias de adenovirus (excepto que aquí el serotipo de referencia es SEQ ID NO:2).

Opcionalmente, dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 40 funcional, y el gen que codifica una proteína Rep 40 funcional comprende una secuencia de ácido nucleico que tiene al menos un 95 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad con la longitud completa o con un fragmento de al menos 600, al menos 700, al menos 800 o al menos 900 nucleótidos de longitud de un tramo de nucleótidos correspondiente a los nucleótidos 993-2252 menos los nucleótidos 1907-2227 de SEQ ID NO:1, o con tramos de nucleótidos correspondientes en un serotipo diferente de AAV. Por tanto, dicho gen que codifica Rep 40 tiene al menos la identidad especificada anteriormente con un tramo de nucleótidos teórico que consiste en los nucleótidos 993-1906 de SEQ ID NO:1 inmediatamente yuxtapuestos con los nucleótidos 2228-2252 de SEQ ID NO:1 (5'-[993-1906]-[2228-2252]-3') o con un tramo de nucleótidos teórico de un serotipo de AAV diferente.

Opcionalmente, dicho al menos un gen rep de AAV comprende al menos un gen que codifica una proteína Rep 40 funcional, y el gen que codifica una proteína Rep 40 funcional comprende una secuencia de ácido nucleico que tiene al menos un 95 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad con la longitud completa o con un fragmento de al menos 900, al menos 1000, al menos 1100 o al menos 1200 nucleótidos de longitud de los nucleótidos 993-2252 de SEQ ID NO:1, o con un tramo de nucleótidos correspondiente en un serotipo diferente de AAV.

Opcionalmente, dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 68 funcional, y el gen que codifica una proteína Rep 68 funcional comprende una secuencia de ácido nucleico que tiene al menos un 95 %, al menos un 98 %, al menos un 99 %, o un 100 % de identidad con la longitud completa o con un fragmento de al menos 1000, al menos 1400, al menos 1500 o al menos 1600 nucleótidos de longitud de un tramo de nucleótidos correspondiente a los nucleótidos 321-2252 menos los nucleótidos 1907-2227 de SEQ ID NO:1, o con tramos de nucleótidos correspondientes en un serotipo diferente de AAV. Por tanto, tal gen que codifica Rep 68 tiene al menos la identidad especificada anteriormente con un tramo de nucleótidos teórico que consiste en los nucleótidos 321-1906 de SEQ ID NO:1 inmediatamente yuxtapuestos con los nucleótidos 2228-2252 de SEQ ID NO:1 (5'- [321-1906]-[2228-2252]-3'), o con un tramo de nucleótidos teórico de un serotipo de AAV diferente.

Dado que Rep 68 no comprende ningún aminoácido codificado por el intrón (nucleótidos 1907-2227), el gen que codifica una proteína Rep 68 funcional no necesita comprender los nucleótidos correspondientes a los nucleótidos 1907-2227. De hecho, la exclusión de los nucleótidos 1907-2227 de dicho al menos un gen rep de AAV asegura que Rep 78 no se codifique (ya que Rep 78 comprende aminoácidos codificados por el intrón). Preferentemente, el plásmido auxiliar no comprende un gen que codifica una proteína Rep 78 funcional.

Opcionalmente, dicho al menos un gen rep de AAV comprende un gen que codifica las proteínas Rep 68 y Rep 40 funcionales, en el que dicho gen comprende un ácido nucleico que tiene al menos un 95 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad con la longitud completa o con un fragmento de al menos 1400, 1500, 1600 o 1700 nucleótidos de longitud de los siguientes tramos de la secuencia de AAV2 nativa (SEQ ID NO:1) colocados en yuxtaposición inmediata desde 5' a 3': 200-1906; 2228-2309, o con tramos yuxtapuestos de nucleótidos correspondientes de un serotipo diferente de AAV.

Opcionalmente, dicho al menos un gen rep de AAV comprende un gen que codifica las proteínas Rep 52 y Rep 40 funcionales, en el que dicho gen comprende un ácido nucleico que tiene al menos un 95 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad con la longitud completa o con un fragmento de al menos 1300, 1400, 1500 o 1600 nucleótidos de longitud del siguiente tramo de la secuencia de AAV2 nativa (SEQ ID NO:1): 658-2300, o con un tramo de nucleótidos correspondiente de un serotipo diferente de AAV.

Opcionalmente, dicho al menos un gen rep de AAV comprende un tramo de nucleótidos que codifican las proteínas Rep 68, Rep 52 y Rep 40 funcionales, en el que dicho tramo comprende un ácido nucleico que tiene al menos un 95 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad con la longitud completa o con un fragmento de al menos 3000, 3200, 3300 o 3400 nucleótidos de longitud de los siguientes tramos de la secuencia de AAV2 nativa (SEQ ID NO:1) colocados en yuxtaposición inmediata desde 5' a 3': 200-1906; 2228-2309; 658-2300, o con tramos yuxtapuestos de nucleótidos correspondientes de un serotipo diferente de AAV.

Opcionalmente, el sistema de dos plásmidos o plásmido auxiliar no comprende una secuencia contigua de al menos 1700, al menos 1800, o 1866 nucleótidos correspondientes a un tramo de nucleótidos contiguo de longitud equivalente comprendido dentro de los nucleótidos 321-2186 de SEQ ID NO:1, o dentro de un tramo de nucleótidos correspondiente en un serotipo diferente de AAV. El tramo de nucleótidos contiguo comprendido dentro de los nucleótidos 321-2186 corresponde a Rep 78.

En algunas realizaciones, dicho al menos un gen rep no comprende un promotor p40 interno funcional. El gen rep/cap nativo comprende un promotor p40 (D.J. Pereira y N. Muzyczka (1997), J. Virol., 71:1747-1756). El promotor p40 dirige la expresión del gen cap, pero no es necesario para la expresión de los genes rep. Un promotor p40 funcional es aquel que es capaz de dirigir la expresión del gen cap.

Puede determinarse si un promotor p40 dado es funcional o no ensayando su capacidad para dirigir la expresión de una proteína (usando un ensayo de expresión). Por ejemplo, el usuario puede preparar un vector de "referencia" que comprende un promotor p40 nativo (como el de los nucleótidos 1710-1827 de SEQ ID NO:1) secuencia arriba de una proteína indicadora, como GFP. A continuación, el usuario puede preparar un vector de "ensayo" que es idéntico al vector de referencia, excepto que el promotor p40 nativo se reemplaza por el promotor p40 cuya funcionalidad se va a analizar (el promotor p40 de "ensayo"). Los dos vectores pueden transfectarse en células huésped adecuadas y las células huésped pueden incubarse en condiciones adecuadas para la expresión de la proteína indicadora. El nivel de proteína indicadora expresada en las células huésped podría medirse, por ejemplo, mediante espectroscopía de fluorescencia. El nivel de expresión del vector de referencia (correspondiente al nivel de expresión dirigido por el promotor p40 nativo) puede compararse luego con el nivel de expresión del vector de ensayo (correspondiente al nivel de expresión dirigido por el promotor de ensayo).

Se considerará que un promotor p40 es funcional si dirige la expresión a un nivel de al menos el 40 % de la expresión dirigida por el promotor p40 nativo. Opcionalmente, un promotor funcional de p40 dirige la expresión a un nivel de al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 %, o al menos el 95 % de la expresión dirigida por el promotor p40 nativo. Opcionalmente, dicho al menos un gen rep de AAV no comprende un

promotor p40 interno funcional si no comprende una secuencia correspondiente al promotor p40 (por ejemplo, los nucleótidos 1710-1827 de SEQ ID NO:1) que dirige la expresión a un nivel de al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 %, o al menos el 95 % de la expresión dirigida por el promotor p40 nativo.

Los promotores P40 que carecen de "caja TATA" no son funcionales. El promotor p40 comprende una caja TATA que comienza en una posición correspondiente a la posición 1823 de SEQ ID NO:1. Por tanto, el plásmido auxiliar que no comprende un promotor p40 interno funcional puede comprender uno o más promotores p40 en los que las cajas TATA están mutadas para hacer que el promotor p40 no sea funcional. Opcionalmente, dicho al menos un gen rep de AAV no comprende un nucleótido T en una posición correspondiente a la posición 1823 de SEQ ID NO:1. Opcionalmente, dicho al menos un gen rep de AAV comprende un nucleótido C en una posición correspondiente a la posición 1823 de SEQ ID NO:1. Opcionalmente, dicho al menos un gen rep de AAV no comprende AAG en las posiciones correspondientes a las posiciones 1826-1828 de SEQ ID NO:1. Opcionalmente, dicho al menos un gen rep de AAV comprende CTC en las posiciones correspondientes a las posiciones 1826-1828 de SEQ ID NO:1.

Ausencia de un gen cap

Un plásmido auxiliar de la invención no comprende un gen cap que codifique un conjunto funcional de proteínas Cap.

En el AAV nativo (tal como un AAV que tiene un genoma como se expone en SEQ ID NO:1), los genes cap de AAV y rep de AAV se solapan. Concretamente, el promotor p40, que dirige la expresión del gen cap de AAV, está presente dentro del gen rep de AAV. Por tanto, cualquier plásmido auxiliar que codifique proteínas Rep de AAV comprenderá nucleótidos del gen cap de AAV. Sin embargo, un plásmido auxiliar de la invención no comprende un tramo de nucleótidos que codifica un conjunto funcional de proteínas Cap. Concretamente, un plásmido auxiliar de la invención puede no comprender un tramo significativo de la secuencia de nucleótidos que sea exclusivamente la secuencia del gen cap de AAV. La expresión "secuencia del gen cap exclusivamente" pretende referirse a la secuencia del gen que codifica una porción de una proteína Cap, pero que no codifica una porción de una proteína Rep de AAV, por ejemplo, los nucleótidos 2253-4410 de SEQ ID NO:1.

Las proteínas Cap de AAV (VP1, VP2 y VP3; ver figura 1) son proteínas que se ensamblan para formar una cápside que rodea el genoma viral. Por tanto, un gen que codifica un conjunto funcional de proteínas Cap de AAV es un gen que codifica proteínas Cap de AAV capaces de ensamblarse para encapsidar un genoma viral.

A conjunto "funcional" de proteínas Cap de AAV es un conjunto que es suficiente para la encapsidación de AAV. Está dentro de las capacidades de los expertos determinar si el producto de un gen Cap de AAV es funcional. Los expertos simplemente necesitan determinar si las proteínas Cap de AAV codificadas facilitan la producción de AAV usando un ensayo de producción de AAV como se describió anteriormente. En este caso, el sistema de dos plásmidos de ensayo comprenderá un vector plasmídico que comprende el gen cap de AAV que codifica las proteínas cuya "funcionalidad" debe determinarse y, por lo demás, el sistema de dos plásmidos de ensayo utilizado será idéntico al sistema de dos plásmidos de referencia. El plásmido auxiliar no comprende un gen cap que codifica un conjunto funcional de proteínas Cap si no comprende ningún material genético correspondiente a un gen cap, o si cualquier gen cap presente en el plásmido auxiliar codifica proteínas que facilitan la producción de AAV en un nivel por debajo del 10 % del nivel facilitado por el producto del gen cap de tipo salvaje, es decir, si el rendimiento de rAAV producido es inferior al 10 % del rendimiento de rAAV producido utilizando el sistema de dos plásmidos de referencia.

En general, un gen cap funcional tiene más de 250 nucleótidos de longitud. En consecuencia, el plásmido auxiliar, que no comprende un gen cap que codifica un conjunto funcional de proteínas Cap, no puede comprender un tramo contiguo de la secuencia del gen cap exclusivamente de más de 250 nucleótidos, de más de 100 nucleótidos, o de más de 60 nucleótidos. Como se ha analizado anteriormente, el gen cap de AAV y el gen rep de AAV se solapan en el AAV nativo. Opcionalmente, el plásmido auxiliar no comprende un tramo contiguo de la secuencia del gen cap exclusivamente de más de 60 nucleótidos. Opcionalmente, el plásmido auxiliar no comprende un gen cap que codifica una proteína VP1 funcional. Opcionalmente, el plásmido auxiliar comprende una parte de la secuencia del gen cap, y la parte de la secuencia del gen cap no codifica un conjunto de proteínas Cap funcionales. El plásmido auxiliar puede comprender una porción de la secuencia del gen cap de AAV, siempre que no codifique un conjunto funcional de proteínas Cap de AAV. Siempre que la parte de la secuencia del gen cap de AAV no sea funcional, se requerirán múltiples acontecimientos de recombinación para proporcionar un rcAAV.

Al menos un gen de virus auxiliar

El plásmido auxiliar comprende al menos un gen de virus auxiliar. El AAV solo puede propagarse en presencia de un virus auxiliar. Los ejemplos de virus auxiliares incluyen adenovirus y virus del herpes.

Sin embargo, como se ha expuesto anteriormente, el cultivo de AAV en presencia de un virus auxiliar no es ventajoso, ya que los virus auxiliares pueden ser líticos para las células, incluidas las células huésped utilizadas para cultivar AAV. Además, si se utilizan virus auxiliares en la producción de productos de rAAV, tal como vectores de terapia génica, el virus auxiliar puede contaminar el producto. Como alternativa a la coinfección con virus auxiliares como el adenovirus, los genes necesarios del virus auxiliar pueden proporcionarse en un plásmido transfectado. Para las células huésped que expresan los genes adenovirales E1A/B (como las células HEK293T), los genes auxiliares necesarios restantes son E4, E2A y VA ARN I y II (un ácido nucleico VA). Si bien estos genes están distribuidos a lo largo de un largo tramo del genoma adenoviral, los presentes inventores han determinado que pueden eliminarse grandes tramos de nucleótidos no codificantes que separan los genes genómicos sin afectar a la expresión de los genes. Por lo tanto, los inventores han diseñado una región mínima de gen auxiliar que es SEQ ID NO:4. SEQ ID NO:4 contiene el ácido nucleico VA y el complemento inverso (como está presente en un plásmido bicatenario) de un gen E2A y un gen E4. El uso de una región mínima de gen auxiliar de este tipo en un plásmido para la producción de rAAV es ventajoso, ya que permite al usuario utilizar un plásmido más pequeño que es menos costoso de producir y más fácil de transfectar a la célula.

Los plásmidos auxiliares de la invención comprenden al menos un gen de virus auxiliar. Preferentemente, los plásmidos auxiliares de la invención comprenden suficientes genes auxiliares para permitir la replicación y la encapsidación de AAV. Puede evaluarse si un plásmido auxiliar comprende o no suficientes genes auxiliares para facilitar la producción de AAV usando un ensayo de producción de AAV como se describió anteriormente. En este caso, el sistema de dos plásmidos de ensayo comprenderá un plásmido auxiliar que comprende los genes auxiliares cuya capacidad para facilitar la producción de AAV se va a analizar y, por lo demás, el sistema de dos plásmidos de ensayo utilizado será idéntico al sistema de dos plásmidos de referencia. En una realización, se considerará que los productos del gen auxiliar facilitan la producción de AAV si facilitan la producción de rAAV en un nivel de al menos el 25 %, al menos el 40 %, al menos el 50 %, al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 % o al menos el 95 % del nivel facilitado por genes auxiliares adenovirales que codifican E4, E2A y VA ARN I y II, es decir, si el rendimiento de rAAV producido es de al menos un 25 %, al menos un 40 %, al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % del rendimiento de rAAV producido utilizando el sistema de dos plásmidos de referencia. Preferentemente, se considerará que los productos del gen auxiliar facilitan la producción de AAV si facilitan la producción de rAAV en un nivel de al menos el 70 %, al menos el 80 %, al menos el 90 % o al menos el 95 % del nivel facilitado por genes auxiliares adenovirales que codifican E4, E2A y VA ARN I y II.

Dado que los plásmidos auxiliares codifican los genes necesarios para permitir la producción eficiente de AAV, no se requiere la adición de un virus auxiliar.

Opcionalmente, dicho al menos un gen de virus auxiliar es un gen de adenovirus. El adenovirus es un virus que se sabe que ayuda a la propagación de AAV (Xiao *et al.* (1998), J. Virol., 72:2224-2232). Opcionalmente, dicho al menos un gen de virus auxiliar es un gen de adenovirus 5 o un gen de adenovirus 2. El genoma del adenovirus 5 se expone en SEQ ID NO:2, y el genoma del adenovirus 2 se expone en SEQ ID NO:3. En consecuencia, los genes auxiliares pueden comprender un tramo de nucleótidos presentes en SEQ ID NO:2 o SEQ ID NO:3, o un tramo de nucleótidos correspondiente en otro serotipo de adenovirus.

Los genes auxiliares de los adenovirus codifican E1A, E1B, E4, E2A y VA ARN I y II.

E1A es codificado por los nucleótidos 560-1545 del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2). Los nucleótidos 560-1545 contienen un intrón del nucleótido 1113 al nucleótido 1228. Este intrón no es fundamental, por lo que un gen E1A que comprende los nucleótidos 560-1112 y 1229-1545 de SEQ ID NO:2 codificaría una proteína E1A funcional.

E1B es en realidad dos proteínas E1B 19K y E1B 55K, que trabajan juntas para bloquear la apoptosis en células infectadas con adenovirus. E1B es codificado por los nucleótidos 1714-2244 (E1B 19K) y por los nucleótidos 2019-3509 (E1B 55 K) del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2).

E4 es codificado por varios marcos de lectura abiertos ("open reading frames", ORF) diferentes del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2). E4 ORF 6/7 es codificado por los nucleótidos 32914-34077, que comprenden un intrón entre los nucleótidos 33193 y 33903. Este intrón no es fundamental, por lo que un E4 ORF 6/7 que comprende los nucleótidos 32914-33192 y 33904-34077 de SEQ ID NO:2 es suficiente. E4 34K es codificado por los nucleótidos 33193-34077 del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2). E4 ORF 4 es codificado por los nucleótidos 33998-34342 del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2). E4 ORF 3 es codificado por los nucleótidos 34353-34703 del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2). E4 ORF B es codificado por los nucleótidos 34700 a 35092 del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2). E4 ORF 1 es codificado por los nucleótidos 35140-35526 del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2).

Una proteína E4 funcional puede comprender solo aminoácidos codificados por los ORF 6 y 7, ya que solo los

aminoácidos codificados por los ORF 6 y 7 son necesarios para la actividad. Opcionalmente, por lo tanto, la proteína E4 funcional comprende una secuencia polipeptídica codificada por todos o por una parte significativa de los ORF 6 y 7. Opcionalmente, la proteína E4 funcional no comprende una secuencia polipeptídica codificada por todos o por una parte de los ORF 1-4 y 34K. Sin embargo, los aminoácidos codificados por los ORF 1-3 y 34K mejoran la actividad de la proteína E4, por lo que, en algunas realizaciones, la proteína E4 funcional comprende aminoácidos codificados por los ORF 1-7.

El gen E2 (E2A) es codificado por los nucleótidos 22443-24032 del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2).

El VA ARN I y II es codificado por los nucleótidos 10589-11044 del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2).

Se cree que E1B y E4 mejoran la acumulación de ARNm de AAV, y se cree que el E2A y VA ARN I y II mejoran el corte y empalme y la traducción del ARNm de AAV. E1B, E4 y E2A son proteínas codificadas por genes presentes en el genoma del adenovirus, mientras que el ácido nucleico VA codifica dos transcritos de ARN conocidos como VA ARN I y VA ARN II. Los transcritos en sí son funcionales en la célula y nunca se traducen en secuencias de aminoácidos. Se apreciará, por lo tanto, que el ácido nucleico VA no codifica una proteína, pero sí "codifica" o "se corresponde con" un ARN, es decir, aunque se usa el término "codificar", el ácido nucleico VA es una secuencia de ácido nucleico no traducida.

De los cinco genes de adenovirus, es posible no incluir E1A o E1B en el plásmido auxiliar, ya que algunas líneas de células huésped (como las células HEK293) expresan uno o más de E1A o E1B de forma constitutiva. Opcionalmente, por lo tanto, el plásmido auxiliar no comprende un gen que codifica una proteína E1A/B adenoviral funcional.

En una realización, dicho al menos un gen del virus auxiliar comprende:

- (a) un ácido nucleico VA (asociado a virus) que codifica los VA ARN I y II funcionales;
- (b) un gen E2A que codifica una proteína E2A funcional; y/o
- (c) un gen E4 que codifica una proteína E4 funcional.

Opcionalmente, dicho al menos un gen del virus auxiliar comprende un ácido nucleico VA, un gen E2A y un gen E4.

Un VA ARN I y II, una proteína E2A o una proteína E4 "funcionales" pueden facilitar la producción de AAV. Está dentro de las habilidades de los expertos determinar si un VA ARN I y II, una proteína E2A o una proteína E4 es funcional. Los expertos simplemente necesitan determinar si el VA ARN I y II, la proteína E2A o la proteína E4 facilitan la producción de AAV utilizando un ensayo de producción de AAV como se describió anteriormente. En este caso, el sistema de dos plásmidos de ensayo comprenderá un plásmido auxiliar que comprende el VA ARN I y II, la proteína E2A o la proteína E4 cuya "funcionalidad" debe determinarse y, por lo demás, el sistema de dos plásmidos de ensayo utilizado será idéntico al sistema de dos plásmidos de referencia. En una realización, se considerará que el VA ARN I y II, la proteína E2A o la proteína E4 son "funcionales" si facilitan la producción de rAAV a un nivel de al menos el 25 %, al menos el 40 %, al menos el 50 %, al menos el 70 %, al menos el 80 %, al menos el 90 % o al menos el 95 % del nivel facilitado por el tipo salvaje (por ejemplo, como se encuentra en el adenovirus 5; SEQ ID NO:2) de VA ARN I y II, proteína E2A o proteína E4, es decir, si el rendimiento de rAAV producido es de al menos un 25 %, al menos un 40 %, al menos un 50 %, al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % del rendimiento de rAAV producido utilizando el sistema de dos plásmidos de referencia. Preferentemente, se considerará que la proteína E4 es "funcional" si facilita la producción de rAAV a un nivel de al menos el 70 %, al menos un 80 %, al menos el 90 % o al menos el 95 % del nivel facilitado por la proteína E4 de tipo salvaje.

Opcionalmente, el gen E4 no se encuentra entre el ácido nucleico VA y el gen E2A, es decir, la secuencia del plásmido es tal que la secuencia del gen E4 no aparece en el plásmido entre la secuencia de ácido nucleico VA y la secuencia del gen E2A. Opcionalmente, el gen E2A se encuentra entre el ácido nucleico VA y el gen E4.

El plásmido auxiliar es de bicatenario y cualquiera de los genes comprendidos en el plásmido auxiliar puede estar presente en una orientación "directa" o "inversa". Por ejemplo, el gen E4 puede comprender los nucleótidos 22443-24032 del genoma del adenovirus 5 (por ejemplo, el genoma de SEQ ID NO:2) o puede comprender el complemento inverso de los nucleótidos 22443-24032. Por tanto, en todos los casos de la solicitud en los que se haga referencia a un plásmido "que comprende" una determinada secuencia de ácido nucleico, las referencias deben interpretarse como que abarcan realizaciones en las que el plásmido comprende el complemento inverso de la secuencia de ácido nucleico.

Los presentes inventores han demostrado que eliminar todo o parte del tramo de nucleótidos presentes en las posiciones 10595-10619 en el genoma del adenovirus (por ejemplo, el genoma de SEQ ID NO:2) reduce significativamente (en aproximadamente un 50 %) la actividad del VA ARN I y II codificados por el ácido nucleico VA. Es ventajoso evitar esta reducción de la actividad.

En consecuencia, en una realización, el ácido nucleico VA tiene un nivel de actividad que tiene al menos un 75%, al menos un 80 %, al menos un 90 %, al menos un 95 %, o entre un 95 % y un 100 % de la actividad de un ácido nucleico VA de tipo salvaje del adenovirus 5. Opcionalmente, el nivel de actividad del ácido nucleico VA se determina midiendo el rendimiento de rAAV, por ejemplo, utilizando el ensayo de producción de AAV descrito anteriormente. En una realización, el ácido nucleico VA comprende una secuencia contigua al menos un 95 %, al menos 98 %, o 100 % idéntica a un tramo de al menos 15 nucleótidos, al menos 20 nucleótidos, o 26 nucleótidos de los nucleótidos 10595-10619 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus.

Opcionalmente, el gen E2A está unido operativamente a un promotor que comprende una secuencia contigua al menos un 96 %, al menos 98 %, o 100 % idéntica a un fragmento de al menos 60, al menos 70, al menos 80 o 100 nucleótidos de los nucleótidos 27037-27136 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus.

La expresión del gen E4 está dirigida por un promotor (denominado "promotor E4" en el presente documento), correspondiente a las posiciones 35585-35848 de la SEQ ID NO:2. Los nucleótidos correspondientes a 35793-35848 del genoma del adenovirus 5 (tal como un genoma de SEQ ID NO:2) son necesarios para que el promotor sea completamente activo. Opcionalmente, por lo tanto, el gen E4 está unido operativamente a un promotor E4 que tiene al menos el 50 %, al menos el 70 %, o al menos el 90 % de la actividad de un promotor de tipo salvaje de adenovirus 5. La actividad del promotor E4 puede determinarse analizando su capacidad para dirigir la expresión de una proteína.

Puede determinarse si un promotor E4 dado es capaz o no de dirigir la expresión del gen E4 usando un ensayo de producción de AAV como se describió anteriormente. Específicamente, si el promotor E4 puede dirigir la expresión del gen E4, facilitará la producción de AAV. En este caso, el sistema de dos plásmidos de ensayo comprenderá un vector plasmídico que comprende el gen E4 y un promotor del gen E4 cuya capacidad para facilitar la producción de AAV se va a analizar, y por lo demás, el sistema de dos plásmidos de ensayo utilizado será idéntico al sistema de dos plásmidos de referencia. En una realización, se considerará que el promotor E4 tiene al menos un 25 %, al menos un 40 %, al menos un 50 %, al menos un 70 %, o al menos un 90 % de la actividad de un promotor E4 de tipo salvaje del adenovirus 5 si facilita la producción de rAAV a un nivel de al menos el 25 %, al menos el 40 %, al menos el 50 %, al menos el 60 %, al menos el 70 %, o al menos el 90 % del nivel facilitado por un promotor del gen E4 de tipo salvaje, es decir, si el rendimiento de rAAV producido es de al menos un 50 %, al menos un 70 %, o al menos un 90 % del rendimiento de rAAV producido usando el sistema de dos plásmidos de referencia. Preferentemente, se considerará que el promotor E4 facilita la producción de AAV si facilita la producción de rAAV en un nivel de al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % del rendimiento de rAAV producido utilizando el sistema de dos plásmidos de referencia.

Reducir el tamaño del plásmido auxiliar

Los presentes inventores han determinado que hay una cantidad significativa de secuencia de ácido nucleico de gen no auxiliar entre las diferentes secuencias de genes auxiliares en el adenovirus 5 nativo, y que una parte sustancial de esta secuencia de ácido nucleico se puede eliminar sin afectar significativamente el nivel de expresión y la función de las proteínas codificadas por las secuencias de genes auxiliares. En consecuencia, los inventores han logrado definir una región de gen auxiliar con un tamaño reducido, y esta puede incluirse en el plásmido auxiliar para reducir su tamaño. Tener una región de gen auxiliar de tamaño reducido es ventajoso, ya que permite reducir el tamaño total del plásmido auxiliar. Los plásmidos más pequeños son más baratos de producir y más fáciles de introducir en las células huésped.

En consecuencia, el plásmido auxiliar puede tener menos de 25000 pb, menos de 20000 pb, menos de 15000 pb, menos de 14500 pb, entre 10000 pb y 25000 pb, entre 10000 pb y 20000 pb, entre 12000 pb y 15000 pb, o aproximadamente 14021 pb de longitud.

De forma análoga, dicho al menos un gen del virus auxiliar comprende un ácido nucleico VA, un gen E2A y un gen E4, y el ácido nucleico VA, el gen E2A y el gen E4 están comprendidos dentro de un tramo de nucleótidos contiguo en el plásmido auxiliar de menos de 15000, menos de 12000, menos de 10000, menos de 9000 o menos de 8500 nucleótidos.

En SEQ ID NO:4 se expone un ejemplo de una región de gen auxiliar con un tamaño reducido. La región de gen auxiliar es una región contigua del plásmido que comienza en el primer nucleótido que codifica un gen auxiliar y termina en el último nucleótido que codifica un gen auxiliar. Opcionalmente, la región de gen auxiliar comprende una secuencia que tiene al menos un 95 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad con la longitud completa o con un fragmento de al menos 6000 nucleótidos, al menos 7000 o al menos 8000 nucleótidos de longitud de SEQ ID NO:4, es decir, dicho al menos un gen del virus auxiliar está comprendido en un tramo contiguo del plásmido que tiene al menos un 95 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad con la longitud completa o con un fragmento de al menos 6000 nucleótidos, al menos 7000 nucleótidos, o al menos 8000

nucleótidos de longitud de SEQ ID NO:4. Opcionalmente, la región de gen auxiliar consiste en una secuencia que tiene al menos un 95 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad con la longitud completa o con un fragmento de al menos 6000 nucleótidos, al menos 7000 o al menos 8000 nucleótidos de longitud de SEQ ID NO:4, es decir, dicho al menos un gen del virus auxiliar está comprendido en un tramo contiguo del plásmido que

Los presentes inventores han identificado regiones específicas del genoma del adenovirus 5 que pueden excluirse del plásmido auxiliar.

En una realización, el plásmido auxiliar no comprende una secuencia contigua de al menos 2000, al menos 2500, al menos 3000, o 3427 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendido dentro de los nucleótidos 194-3620 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus. En una realización, el plásmido auxiliar no comprende una secuencia contigua de al menos 50, al menos 60, o 69 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendido dentro de los nucleótidos 4032-4100 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus.

En una realización, el plásmido auxiliar no comprende una secuencia contigua de al menos 15000, al menos 20000, al menos 22000, o 22137 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendido dentro de los nucleótidos 10619-32755 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus.

En una realización, el plásmido auxiliar no comprende una secuencia contigua de al menos 15000, al menos 20000, al menos 21000, o 21711 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendido dentro de los nucleótidos 11045-32755 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus.

De forma análoga, los inventores han identificado regiones del genoma AAV2 que no necesitan incluirse en el plásmido auxiliar y, en consecuencia, pueden excluirse con el fin de minimizar el tamaño del plásmido auxiliar. En consecuencia, en una realización, el plásmido auxiliar no comprende una secuencia contigua de al menos 200, al menos 300, al menos 350, o 363 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendido dentro de los nucleótidos 4051-4413 de SEQ ID NO:1, o un tramo de nucleótidos correspondiente en un serotipo diferente de AAV. De forma análoga, en una realización, el plásmido auxiliar no comprende una secuencia contigua de al menos 400, al menos 500, al menos 600, o 647 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendido dentro de los nucleótidos 2301-2947 de SEQ ID NO:1, o un tramo de nucleótidos correspondiente en un serotipo diferente de AAV.

Sitios de unión a rep artificial

Las proteínas rep se unen a un sitio de unión a rep ("rep binding site", RBS), que es una secuencia de ácido nucleico que tiene una secuencia de consenso de GCTCGCTCGCTCGCTC (McCarty, D.M. *et al.* (1994), J. Virol., 68(8): 4988-4997)). En consecuencia, un sitio de unión a rep es una secuencia de ácido nucleico que tiene homología con GCTCGCTCGCTCGCTC. Se puede determinar si una secuencia de ácido nucleico (secuencia de ácido nucleico de ensayo) comprende un sitio de unión a rep, por ejemplo, determinar si una secuencia de ácido nucleico comprende una secuencia que tiene homología con GCTCGCTCGCTCGCTC, utilizando un ensayo de desplazamiento de movilidad electroforética ("electrophoretic mobility shift assay", EMSA). La secuencia de ácido nucleico de ensayo se aplica a un primer pocillo de un gel adecuado para usar en electroforesis, y una mezcla de la secuencia de ácido nucleico de ensayo y las proteínas Rep68/Rep78 se aplica a un segundo pocillo del gel. Si la secuencia de ácido nucleico de ensayo comprende un sitio de unión a rep, esto provocará un desplazamiento en la movilidad de la secuencia de ácido nucleico en el pocillo que comprende las proteínas Rep68/Rep78 en comparación con el pocillo que carece de las proteínas Rep68/Rep78.

La unión de proteínas Rep a RBS aberrantes podría dar como resultado una recombinación no homóloga mediada por Rep entre diferentes secuencias que comprenden RBS, dando como resultado secuencias de fusión no deseadas. Por ejemplo, las proteínas Rep podrían unirse al RBS en las ITR de AAV y cualquier RBS presente en un esqueleto de plásmido bacteriano, lo que da como resultado la fusión de las secuencias del esqueleto del plásmido en una ITR. Esto podría dar como resultado una replicación y una encapsidación mejoradas de dichas secuencias de fusión no deseadas.

En consecuencia, se prefiere que los únicos RBS presentes en los plásmidos de la invención correspondan a los RBS presentes en AAV de tipo salvaje (lo cual permite que las proteínas Rep realicen su función normal). Sin embargo, cuando se construye un plásmido, es posible que esté presente un RBS, ya sea en el esqueleto del plásmido o en cualquier otra parte del plásmido en secuencias no derivadas de AAV, y dichas secuencias se consideran RBS "artificiales" o "aberrantes" y deben evitarse/eliminarse.

Una fuente común de RBS artificiales en plásmidos es del esqueleto del plásmido. El esqueleto del plásmido es la secuencia bacteriana necesaria para amplificar los plásmidos en las células huésped bacterianas. El esqueleto del plásmido puede ser la región del plásmido que no se deriva de adenovirus o AAV o que no está situada entre dos ITR derivadas de AAV. Opcionalmente, el esqueleto del vector plasmídico abarca cualquier nucleótido excepto el gen cap, un promotor (región) unido operativamente al gen cap, ITR y el módulo de expresión. Opcionalmente, el esqueleto del plásmido auxiliar abarca cualquier nucleótido excepto al menos un gen auxiliar y los elementos reguladores asociados derivados del adenovirus, dicho al menos un gen rep y los elementos reguladores asociados derivados de AAV, y uno o más promotores unidos operativamente a dicho al menos un gen rep. En una realización, el plásmido auxiliar y/o el vector plasmídico comprende un esqueleto, y el esqueleto plasmídico no comprende un RBS artificial. Opcionalmente, el plásmido auxiliar y el vector plasmídico comprenden cada uno un esqueleto, y ninguno de los esqueletos plasmídicos comprende un RBS artificial.

En consecuencia, en una realización, el plásmido auxiliar y/o el vector plasmídico no comprende un RBS artificial. Preferentemente, el plásmido auxiliar y el vector plasmídico no comprenden un RBS artificial. Opcionalmente, no hay RBS en el vector plasmídico y/o en el plásmido auxiliar, excepto dentro de cualquier ITR, promotor p5 y promotor p19.

Gen cap de AAV

El vector plasmídico puede comprender un gen cap de AAV. El gen cap de AAV codifica una proteína Cap de AAV funcional. El gen cap de AAV puede codificar un conjunto funcional de proteínas Cap de AAV. El AAV generalmente comprende tres proteínas Cap de AAV, VP1, VP2 y VP3. Estas tres proteínas forman una cápside en la que se inserta el genoma de AAV y permiten la transferencia del genoma de AAV a una célula huésped. Todas las VP1, VP2 y VP3 son codificados en los AAV nativos por un solo gen, el gen cap de AAV. La secuencia de aminoácidos de VP1 comprende la secuencia de VP2. La parte de VP1 que no forma parte de VP2 se denomina VP1unique o VP1U. La secuencia de aminoácidos de VP2 comprende la secuencia de VP3. La parte de VP2 que no forma parte de VP3 se denomina VP2unique o VP2U.

A conjunto "funcional" de proteínas Cap de AAV es aquel que permite la encapsidación de AAV. Como se ha analizado anteriormente, está dentro de las capacidades de los expertos determinar si una proteína Cap de AAV dada o un conjunto de proteínas Cap de AAV son funcionales. Los expertos simplemente necesitan determinar si las proteínas Cap de AAV codificadas facilitan la producción de AAV usando un ensayo de producción de AAV como se describió anteriormente. En este caso, el sistema de dos plásmidos de ensayo comprenderá un vector plasmídico que comprende el gen cap de AAV que codifica las proteínas Cap de AAV cuya "funcionalidad" debe determinarse y, por lo demás, el sistema de dos plásmidos de ensayo utilizado será idéntico al sistema de dos plásmidos de referencia. Las proteínas Cap de AAV se considerarán "funcionales" si facilitan la producción de rAAV a un nivel de al menos el 25 %, al menos un 40 %, al menos un 50 %, al menos un 70 %, al menos un 80 %, al menos el 90 % o al menos el 95 % del nivel facilitado por el producto del gen cap de AAV de tipo salvaje, es decir, si el rendimiento de rAAV producido es de al menos un 25 %, al menos un 40 %, al menos un 50 %, al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % del rendimiento de rAAV producido utilizando el sistema de dos plásmidos de referencia. Preferentemente, las proteínas Cap de AAV se considerarán "funcionales" si facilitan la producción de rAAV a un nivel de al menos el 70 %, al menos el 80 %, al menos el 90 % o al menos el 95 % del nivel facilitado por el producto del gen cap de AAV de tipo salvaje.

Opcionalmente, las proteínas VP2 y/o VP3 son "funcionales" si un AAV que comprende las proteínas VP2 y/o VP3 es capaz de transducir células Huh7 a un nivel de al menos un 25 %, al menos un 40 %, al menos un 50 %, al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % de el de un AAV equivalente que comprende una proteína VP2 y/o VP3 de tipo salvaje. La capacidad de una partícula de AAV para transducir células Huh7 se puede analizar agregando una proteína indicadora, como la proteína fluorescente verde ("green fluorescent protein", GFP), a la partícula AAV, mezclando la partícula de AAV con células Huh7 y midiendo la fluorescencia producida.

Opcionalmente, el vector plasmídico comprende un gen cap de AAV que codifica una proteína VP1, VP2 y/o VP3. Opcionalmente, las proteínas VP1, VP2 y VP3 se expresan a partir de más de un gen cap de AAV. Opcionalmente, el vector plasmídico comprende un gen cap de AAV que codifica una proteína VP1, VP2 y VP3. Opcionalmente, el vector plasmídico comprende un gen cap de AAV que codifica una VP1 funcional, es decir, una proteína VP1 capaz de ensamblarse con otras proteínas Cap de AAV para encapsidar un genoma viral.

Diferentes serotipos de AAV tienen proteínas Cap de AAV que tienen diferentes secuencias de aminoácidos. Un gen cap de AAV que codifica cualquier proteína (conjunto de proteínas) Cap de AAV es adecuado para su uso en relación con la presente invención. La proteína Cap de AAV puede ser una proteína Cap de AAV nativa expresada en AAV de un determinado serotipo. Como alternativa, la proteína Cap de AAV puede ser no natural, por ejemplo, una proteína Cap de AAV modificada que se diseña para comprender una secuencia diferente a la de una proteína Cap de AAV nativa. Los genes que codifican proteínas Cap no naturales son particularmente ventajosos, puesto que, en el contexto de las aplicaciones de terapia génica, es posible que menos pacientes potenciales tengan

niveles de anticuerpos que impidan la transducción por AAV que comprenden proteínas Cap no naturales, en relación con las cápsides nativas.

Opcionalmente, el gen cap de AAV codifica una proteína Cap de AAV de un serotipo seleccionado del grupo que consiste en los serotipos 1, 2, 3A, 3B, 4, 5, 6, 7, 8, 9, 10, 11, 12 o 13. Opcionalmente, el gen cap de AAV codifica una proteína Cap de AAV seleccionada del grupo que consiste en LK03, rh74, rh10 y Mut C (documento WO 2016/181123; documento WO 2013/029030; documento WO 2017/096164).

Promotor del gen cap de AAV

Opcionalmente, el vector plasmídico comprende un promotor del gen cap de AAV. El promotor del gen cap de AAV puede estar unido operativamente a un gen cap de AAV. Como alternativa, el vector plasmídico puede no comprender un gen cap de AAV, pero puede comprender un sitio de clonación unido operativamente (es decir, en estrecha yuxtaposición: 5'-[promotor del gen cap]-[sitio de clonación]-3') al promotor del gen cap. El sitio de clonación puede ser un sitio de clonación múltiple ("multiple cloning site", MCS o policonector; véanse, por ejemplo, las figuras 4C-E). El usuario puede desear tener la opción de agregar un gen cap específico para una aplicación específica. Por ejemplo, si el vector plasmídico se va a utilizar para producir AAV para su uso en terapia génica, el usuario puede desear que el vector plasmídico carezca de un gen cap, pero que comprenda un sitio de clonación para permitir clonar un gen cap específico para una aplicación específica. El vector plasmídico podría usarse en relación con cualquier transgén (en un módulo de expresión), y el usuario puede descubrir que, para ciertos transgenes, la encapsidación de dichos módulos en cápsides que tienen propiedades, tales como tropismo hepático, es ventajosa, mientras que, para otros transgenes, son ventajosas las cápsides que tienen diferentes tropismos. Mediante el diseño de un vector plasmídico que comprende un promotor del gen cap relacionado con un sitio de clonación, el usuario puede fácilmente "enchufar" un gen cap apropiado para una aplicación específica (como el uso de un transgén específico).

El gen cap de AAV nativo (es decir, el gen cap de un AAV de tipo salvaje) está unido operativamente a un promotor p40, un promotor p5 y un promotor p19. Opcionalmente, dicho al menos un promotor del gen cap de AAV comprende un promotor p40 de AAV, un promotor p5 y/o un promotor p19. Opcionalmente, dicho al menos un promotor del gen cap de AAV comprende un promotor p40 de AAV, un promotor p5 y un promotor p19. Sin embargo, puede usarse cualquier promotor adecuado que sea capaz de dirigir la expresión del gen cap de AAV.

Puede determinarse si un promotor del gen cap de AAV dado es capaz o no de dirigir la expresión del gen cap de AAV usando un ensayo de producción de AAV como se describió anteriormente. Específicamente, si el promotor del gen cap de AAV puede dirigir la expresión del gen cap de AAV, facilitará la producción de AAV. En este caso, el sistema de dos plásmidos de ensayo comprenderá un vector plasmídico que comprende el gen cap de AAV y un promotor del gen cap de AAV cuya capacidad para facilitar la producción de AAV se va a analizar y, por lo demás, el sistema de dos plásmidos de ensayo utilizado será idéntico al sistema de dos plásmidos de referencia. En una realización, se considerará que el promotor del gen cap de AAV facilita la producción de AAV si facilita la producción de rAAV en un nivel de al menos el 25 %, al menos el 40 %, al menos el 50 %, al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 % o al menos el 95 % del nivel facilitado por un promotor p40 del gen cap de tipo salvaje, es decir, si el rendimiento de los vectores rAAV producidos es de al menos un 25 %, al menos un 40 %, al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % del rendimiento de rAAV producido utilizando el sistema de dos plásmidos de referencia. Preferentemente, se considerará que el promotor del gen cap de AAV facilita la producción de AAV si facilita la producción de rAAV en un nivel de al menos el 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % del rendimiento de rAAV producido utilizando el sistema de dos plásmidos de referencia.

Se entiende que el promotor p40 dirige la expresión del gen cap. El promotor p40 nativo está contenido dentro del gen rep nativo. El promotor p40 de AAV2 tiene una secuencia de nucleótidos 1710-1827 de SEQ ID NO:1. Sin embargo, los promotores p40 en otros serotipos de AAV pueden comprender secuencias ligeramente diferentes. En algunas realizaciones, el promotor p40 tiene una secuencia al menos un 95 %, al menos un 98 %, o un 99 % idéntica a los nucleótidos 1710-1827 de SEQ ID NO:1, o una secuencia correspondiente de otro serotipo de AAV. En algunas realizaciones, el promotor p40 tiene una secuencia idéntica en al menos un 98 % a los nucleótidos 1710-1827 de SEQ ID NO:1.

Los presentes inventores han descubierto que al menos un promotor del gen cap de AAV que comprende un promotor p40, un promotor p5 y un promotor p19 es ventajoso, puesto que el promotor p5 y el promotor p19 desempeñan un papel en la regulación del promotor p40, y la presencia del promotor p40, el promotor p5 y el promotor p19 conduce a la expresión oportunamente regulada del gen cap de AAV que da como resultado un rAAV de alto rendimiento/alta calidad. En consecuencia, en algunas realizaciones, dicho al menos un promotor del gen cap de AAV está comprendido en una "región promotora" que comprende un promotor p40, un promotor p5 y un promotor p19. Opcionalmente, la región promotora comprende una secuencia con al menos un 95 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad con la longitud completa o con un fragmento de al menos 800, al menos 900, al menos 1000 o al menos 1100 nucleótidos de longitud de los siguientes tramos de la secuencia nativa de AAV2 (SEQ ID NO:1) colocados en yuxtaposición inmediata de 5' a 3': 200-354; 600-1049; 1701-2202, o con

tramos yuxtapuestos de nucleótidos correspondientes de un serotipo diferente de AAV.

El promotor p5 nativo está secuencia arriba del gen rep de AAV nativo. El promotor p5 de AAV2 tiene una secuencia de nucleótidos 204-292 de SEQ ID NO:1. Sin embargo, los promotores p5 en otros serotipos de AAV pueden comprender secuencias ligeramente diferentes. En algunas realizaciones, el promotor p5 tiene una secuencia al menos un 95 %, al menos un 98 %, o al menos un 99 % idéntica a los nucleótidos 204-292 de SEQ ID NO:1, o una secuencia correspondiente de otro serotipo de AAV. En algunas realizaciones, el promotor p5 tiene una secuencia idéntica en al menos un 98 % a los nucleótidos 204-292 de SEQ ID NO:1.

El promotor p19 está contenido dentro del gen rep de AAV nativo. El promotor p19 de AAV2 tiene una secuencia de nucleótidos 730-890 de SEQ ID NO:1. Sin embargo, los promotores p19 en otros serotipos de AAV pueden comprender secuencias ligeramente diferentes. En algunas realizaciones, el promotor p19 tiene una secuencia al menos un 95 %, al menos un 98 %, o al menos un 99 % idéntica a los nucleótidos 730-890 de SEQ ID NO:1, o una secuencia correspondiente de otro serotipo de AAV. En algunas realizaciones, el promotor p19 tiene una secuencia idéntica en al menos un 98 % a los nucleótidos 730-890 de SEQ ID NO:1.

Módulo de expresión flanqueado en al menos un lado por una ITR

El sistema de dos plásmidos de la invención puede usarse para producir un vector AAV para su uso en terapia génica. La "terapia de genes" implica administrar AAV/partículas virales de la invención que son capaces de expresar un transgén (tal como una secuencia de nucleótidos que codifica el factor IX) en el huésped al que se administran. En estos casos, el vector plasmídico comprenderá un módulo de expresión.

Opcionalmente, el vector plasmídico comprende al menos una ITR. Por tanto, el vector plasmídico comprende al menos una ITR, pero, más normalmente, dos ITR (generalmente con uno de los extremos del módulo de expresión, es decir, una en el extremo 5' y la otra en el extremo 3'). Puede haber secuencias intermedias entre el módulo de expresión y una o más de las ITR. El módulo de expresión puede incorporarse en una partícula viral ubicado entre dos ITR normales o ubicado a ambos lados de una ITR diseñada con dos regiones D. Opcionalmente, el vector plasmídico comprende secuencias ITR que se derivan de AAV1, AAV2, AAV4 y/o AAV6. Preferentemente, las secuencias ITR son secuencias ITR de AAV2.

Opcionalmente, el vector plasmídico comprende un módulo de expresión. Como se describe en el presente documento, un módulo de expresión se refiere a una secuencia de ácidos nucleicos que comprende un transgén y un promotor unido operativamente al transgén. Opcionalmente, el módulo comprende además elementos reguladores de la transcripción adicionales, tales como potenciadores, intrones, regiones no traducidas, terminadores transcripcionales, etc.

Opcionalmente, el módulo de expresión comprende un elemento regulador de la transcripción que comprende el elemento promotor y/o el elemento potenciador de HLP2, HLP1, LP1, HCR-hAAT, ApoE-hAAT y/o LSP. Estos elementos reguladores de la transcripción se describen con más detalle en las siguientes referencias bibliográficas: HLP2: documento WO16/075473; HLP1: McIntosh J. *et al.*, Blood, 25 de abril de 2013, 121(17):3335-3344; LP1: Nathwani *et al.*, Blood. 1 de abril de 2006, 107(7): 2653-2661; HCR-hAAT: Miao *et al.*, Mol. Ther., 2000, 1:522-532; ApoE-hAAT: Okuyama *et al.*, Human Gene Therapy, 7, 637-645 (1996); y LSP: Wang *et al.*, Proc. Natl. Acad. Sci. USA, 1999, 30 de marzo, 96(7): 3906-3910. Cada uno de estos elementos reguladores de la transcripción comprende un promotor, un potenciador y, opcionalmente, otros nucleótidos. Si el polinucleótido está destinado a la expresión en el hígado, el promotor puede ser un promotor específico de hígado. Opcionalmente, el promotor es un promotor específico de hígado humano.

El transgén puede ser cualquier gen adecuado. Si el vector plasmídico es para su uso en terapia génica, el transgén puede ser cualquier gen que comprenda o codifique una proteína o una secuencia de nucleótidos que pueda usarse para tratar una enfermedad. Por ejemplo, el transgén puede codificar una enzima, una proteína metabólica, una proteína de señalización, un anticuerpo, un fragmento de anticuerpo, una proteína similar a un anticuerpo, un antígeno, o un ARN no traducido, como un miARN, ARNip, ARNnp o ARN antisentido.

Codones de iniciación de la traducción prescindibles

Opcionalmente, el vector plasmídico no comprende codones de iniciación de la traducción prescindibles. Opcionalmente, el vector plasmídico puede comprender al menos dos genes que deben poder transcribirse y traducirse (el transgén (en un módulo de expresión) y el gen cap de AAV). Opcionalmente, el transgén codifica un ARN funcional que no codifica una proteína o polipéptido traduccional. El transgén, si codifica una proteína, y el gen cap de AAV comprenderán codones de iniciación (ATG o GTG) para activar la iniciación de la traducción del gen. Sin embargo, el vector plasmídico puede comprender casos adicionales de ATG o GTG (ya sea dentro o fuera del marco con el marco de lectura de estos genes), y es posible que la traducción pueda iniciarse en una de estas posiciones. Dado que no es necesario iniciar la traducción en el sitio de estos casos adicionales de ATG o GTG, los codones ATG o GTG pueden considerarse "codones de iniciación de la traducción prescindibles". Se prefiere que los codones de iniciación de la traducción prescindibles se eliminen o se vuelvan no funcionales, en concreto cuando aparecen

en la región promotora. La región promotora es una región del vector plasmídico que comprende uno o más promotores unidos operativamente al gen cap de AAV. En algunas realizaciones, el gen cap de AAV está unido operativamente a uno o más de los promotores p5, p19 y p40, y en tales realizaciones la región promotora comprende los promotores p5, p19 y p40.

Opcionalmente, el plásmido comprende una región promotora que comprende uno o más promotores, y la región promotora no comprende codones ATG ni GTG. Opcionalmente, la región promotora comprende los promotores p5, p19 y/o p40, y los codones ATG o GTG en una o más posiciones correspondientes a las posiciones (a) 321-323 (codón de iniciación ATG Rep78/68), (b) 766-768 (codón ATG), (c) 955-957 (codón ATG), (d) 993-995 (codón de iniciación ATG Rep52/40) y (e) 1014-1016 (codón GTG) de SEQ ID NO:1 están ausentes o mutados.

Opcionalmente, en la región promotora:

- (a) los nucleótidos correspondientes a los nucleótidos 321-323 de SEQ ID NO:1 están ausentes;
- (b) los nucleótidos correspondientes a los nucleótidos 766-768 de SEQ ID NO:1 no son ATG y son opcionalmente ATT;
- (c) los nucleótidos correspondientes a los nucleótidos 955-957 de SEQ ID NO:1 están ausentes;
- (d) los nucleótidos correspondientes a los nucleótidos 993-995 de SEQ ID NO:1 están ausentes; y/o
- (e) los nucleótidos correspondientes a los nucleótidos 1014-1016 de SEQ ID NO:1 están ausentes.

Opcionalmente, en la región promotora:

- (a) los nucleótidos correspondientes a los nucleótidos 321-323 de SEQ ID NO:1 están ausentes;
- (b) los nucleótidos correspondientes a los nucleótidos 766-768 de SEQ ID NO:1 no son ATG y son opcionalmente ATT;
- (c) los nucleótidos correspondientes a los nucleótidos 955-957 de SEQ ID NO:1 están ausentes;
- (d) los nucleótidos correspondientes a los nucleótidos 993-995 de SEQ ID NO:1 están ausentes; y
- (e) los nucleótidos correspondientes a los nucleótidos 1014-1016 de SEQ ID NO:1 están ausentes.

Tamaño del vector plasmídico

Tal como se analizó anteriormente en relación con el plásmido auxiliar, es ventajoso que los plásmidos usados en la producción de rAAV sean lo más pequeños posible. Los plásmidos más pequeños son más baratos de producir y más fáciles de transfectar a células.

El tamaño del vector plasmídico estará definido, en parte, por el tamaño de los promotores, los elementos reguladores de la transcripción, los transgenes y los genes cap de AAV que estén incluidos y, tal como se analizó anteriormente, la naturaleza de estas características estará definida parcial o totalmente por la aplicación para la que se utilizará el rAAV.

Sin embargo, el tamaño del vector plasmídico se puede reducir asegurándose de que el tamaño del esqueleto que se utiliza sea pequeño y/o asegurándose de que el vector plasmídico no contenga un espaciador (regiones contiguas de secuencias de ácido nucleico no codificantes de más de 200 nucleótidos de longitud).

Opcionalmente, el vector plasmídico comprende un esqueleto de menos de 4000 nucleótidos, menos de 3500 nucleótidos, menos de 3000 nucleótidos, o menos de 2500 nucleótidos de longitud. Opcionalmente, el vector plasmídico no comprende ningún espaciador.

Célula huésped

Se describe una célula huésped que comprende el sistema de dos plásmidos de la invención, el plásmido auxiliar de la invención o el vector plasmídico.

La célula huésped es preferentemente una célula huésped que puede usarse para producir rAAV. Por lo tanto, la célula huésped puede ser una célula huésped adecuada para la producción de rAAV. Además, la célula huésped puede ser para la producción de rAAV.

En general, una célula huésped que es adecuada para la producción de rAAV es una célula huésped que procede de una línea celular eucariota, preferentemente una línea celular de vertebrados, preferentemente una línea celular de mamífero, preferentemente una línea celular humana. Opcionalmente, la célula huésped es una célula que se selecciona del grupo que consiste en una célula HEK293T, una célula HEK293, una célula HEK293EBNA, una célula CAP, una célula CAP-T, una célula AGE1.CR, una célula PerC6, una célula C139, una célula EB66, una célula BHK, una célula COS, una célula Vero, una célula Hela y una célula A549. Preferentemente, la célula huésped se selecciona del grupo que consiste en una célula HEK293T, una célula HEK293, una célula HEK293EBNA, una célula CAP, una célula CAP-T, una célula AGE1.CR, una célula PerC6, una célula C139 y una célula EB66. Incluso más preferentemente, la célula huésped se selecciona del grupo que consiste en una célula HEK293T, una célula

HEK293 y una célula HEK293EBNA. Por ejemplo, la célula huésped puede ser una célula HEK293T. Opcionalmente, la célula huésped es una célula que expresa una proteína E1A/B adenoviral funcional. Por ejemplo, la célula huésped puede comprender un cromosoma que comprende un gen que codifica una proteína E1A/B adenoviral funcional. Las proteínas E1A/B adenovirales funcionales se analizan en la sección titulada "Al menos un gen de virus auxiliar".

Está dentro de las capacidades de los expertos determinar si una célula huésped es adecuada para la producción de rAAV. Los expertos simplemente necesitan determinar si la célula huésped facilita la producción de AAV usando un ensayo de producción de AAV como se describió anteriormente. En este caso, el sistema de dos plásmidos de ensayo y el sistema de dos plásmidos de referencia que se utilizan serán idénticos. Sin embargo, el sistema de dos plásmidos de ensayo se transfectará a la célula huésped cuya idoneidad para la producción de AAV recombinante se va a analizar, y el sistema de dos plásmidos de referencia se transfectará en células HEK293T. La célula huésped se considerará adecuada para la producción de AAV recombinante si facilita la producción de AAV en un nivel de al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 70 %, al menos el 80 %, al menos el 90 % o al menos el 95 % del nivel facilitado por las células HEK293T, es decir, si el rendimiento de AAV recombinante producido es de al menos un 30 %, al menos un 40 %, al menos un 50 %, al menos un 70 %, al menos un 80 %, al menos un 90 % o al menos un 95 % del rendimiento de AAV recombinante producido en células HEK293T.

Nivel bajo de AAV competente para la replicación (rcAAV)

El sistema de dos plásmidos de la invención, el plásmido auxiliar de la invención o el vector plasmídico pueden usarse para producir una preparación de rAAV que tiene un nivel bajo de AAV competente para la replicación (rcAAV).

El plásmido auxiliar o el vector plasmídico también pueden usarse para reducir o minimizar el nivel de rcAAV producido durante la producción de rAAV.

La presente invención también proporciona un método para producir una preparación de rAAV que comprende:

- (a) obtener el sistema de dos plásmidos de la invención, el plásmido auxiliar de la invención o el vector plasmídico;
- (b) transfectar una célula huésped con el sistema de dos plásmidos de la invención, el plásmido auxiliar de la invención o el vector plasmídico; y
- (c) cultivar la célula huésped en condiciones adecuadas para la producción de rAAV

en el que la preparación de rAAV comprende un nivel bajo de AAV competente para la replicación (rcAAV).

También se proporciona un método para reducir o minimizar el nivel de AAV competente para la replicación (rcAAV) producido durante la producción de AAV recombinante (rAAV) que comprende:

- (a) obtener el sistema de dos plásmidos de la invención, el plásmido auxiliar de la invención o el vector plasmídico;
- (b) transfectar una célula huésped con el sistema de dos plásmidos de la invención, el plásmido auxiliar de la invención o el vector plasmídico; y
- (c) cultivar la célula huésped en condiciones adecuadas para la producción de rAAV.

El método puede comprender además una etapa de recolección del rAAV para proporcionar una preparación de rAAV, que opcionalmente tiene un nivel bajo de rcAAV.

"rAAV" o "AAV recombinante" se refiere a partículas de AAV, es decir, partículas que comprenden un genoma de AAV (tal como un genoma de vector) y una cápside de AAV.

Para los fines de la presente invención, el término "rcAAV" o la expresión "AAV competente para la replicación" pretenden referirse a partículas de rAAV que comprenden un genoma que comprende un gen rep o un genoma que comprende un gen cap. Aunque, para los fines de la presente invención, el término "rcAAV" se refiere a partículas de rAAV que comprenden un genoma que comprende un gen rep o un gen cap, se entiende que las partículas de rAAV que comprenden un genoma que comprende un gen rep, pero no un gen cap (denominadas en el presente documento "rcAAV deficiente en cap") o un genoma que comprende un gen cap, pero no un gen rep (denominadas en el presente documento "rcAAV deficiente en rep") no son competentes para la replicación de forma aislada. Más bien, para replicarse (en presencia de las funciones de virus auxiliares necesarias, pero sin ningún requisito para las funciones de AAV proporcionadas *in trans*), una partícula de AAV debe comprender un gen rep y un gen cap (los llamados "rcAAV de tipo pseudosalvaje"). Sin embargo, una partícula de AAV que comprende un gen rep, pero no comprende un gen cap, puede replicarse cuando se coinfecta en una célula huésped con una partícula de AAV que comprende un gen cap, y una partícula de AAV que comprende un gen cap, pero no comprende un gen rep, puede replicarse cuando se coinfecta en una célula huésped con una partícula de AAV que comprende un gen rep. Por tanto, el término "rcAAV" engloba "rcAAV deficiente en cap", pero también "rcAAV de tipo pseudosalvaje" y "rcAAV

deficiente en rep"). El rcAAV de tipo pseudosalvaje puede (en presencia de funciones de virus auxiliares) ser capaz de replicarse en una célula huésped sin coinfección con otra partícula de AAV y producir virus de progenie. El rcAAV de tipo pseudosalvaje puede comprender un gen rep y un gen cap flanqueado por secuencias ITR.

- 5 Todas las especies de rcAAV son indeseables en una preparación de rAAV. El rcAAV deficiente en cap puede replicarse cuando se coinfecta con partículas de AAV que contienen cap. El rcAAV deficiente en rep puede replicarse cuando se coinfecta con partículas de AAV que contienen rep, y el rcAAV de tipo pseudosalvaje puede replicarse sin coinfección con AAV. Si el rcAAV se replica, puede causar efectos secundarios significativos. En consecuencia, es ventajoso reducir o minimizar el nivel de rcAAV producido. Como se usa en el presente documento, la expresión "reducir el nivel de rcAAV" se refiere a disminuir el número de rcAAV que se producen. Tal como se utiliza en el presente documento, la expresión "minimizar el nivel de rcAAV" se refiere a reducir el nivel de rcAAV al nivel más bajo posible dadas las limitaciones del método.

- 15 Cuando se usan los métodos de la invención, el genoma de rAAV que se encapsida no debe comprender un gen rep ni un gen cap, ya que, en general, ni el gen rep ni el gen cap están relacionados (en el mismo plásmido, muy cerca) con una ITR. Sin embargo, los acontecimientos de recombinación únicos podrían dar como resultado la generación de rcAAV deficientes en cap o rep. Para generar rcAAV de tipo pseudosalvaje, deben producirse dos acontecimientos de recombinación, y esto es muy poco probable.

- 20 Puede determinarse si una partícula de rAAV comprende o no un genoma que comprende un gen rep, o si una preparación de partículas de rAAV comprende una partícula que comprende un genoma que comprende un gen rep, usando qPCR. Opcionalmente, la cantidad de rcAAV se mide determinando el número de copias de rep producido. Opcionalmente, la cantidad de rcAAV se mide determinando el número de copias de rep por célula (es decir, célula huésped) usando qPCR.

- 25 Se debe utilizar un cebador (o dos cebadores) específico para el gen rep en la qPCR. Opcionalmente, los cebadores son específicos (inverso y complementario o idéntico dependiendo de si el cebador es un cebador directo o un cebador inverso) para una región de al menos 12, al menos 14, al menos 16, o al menos 18 de los nucleótidos 321-2252 de SEQ ID NO:1, que codifican las cuatro proteínas Rep. Opcionalmente, el cebador es específico para rep 40. Opcionalmente, el cebador es específico para rep 52. Opcionalmente, el cebador es específico para rep 68.

- 35 Para determinar el número de copias de rep por célula usando qPCR, se debe usar un segundo par de cebadores en la qPCR que sean específicos para un gen endógeno (como un gen constitutivo endógeno) en la célula huésped, como la albúmina humana en las células HEK293. Opcionalmente, cada cebador o par de cebadores es específico para una región (complementario con esta) de al menos 8, al menos 10, al menos 12 o al menos 15 nucleótidos de la secuencia genómica de la albúmina humana. El número de copias por célula puede entonces determinarse como la proporción entre el número de copias de rep y el número de copias del gen endógeno.

- 40 Como alternativa, el método anterior para detectar/medir rcAAV (de tipo pseudosalvaje y/o rcAAV deficiente en cap) puede realizarse mediante qPCR específica del gen cap que puede usarse para detectar rcAAV deficiente en rep.

- 45 Los métodos o usos pueden proporcionar una preparación de rAAV. En tales realizaciones, la preparación de rAAV preferentemente comprende un nivel bajo de rcAAV. Un nivel bajo de rcAAV es generalmente un nivel de rcAAV menor que 1 rcAAV en 10^7 rAAV, o es más bajo que el de una preparación de rAAV que se produce utilizando un método equivalente al de un método de la invención, excepto que el vector plasmídico comprende al menos un gen rep y al menos un gen cap. El nivel de rAAV se puede determinar usando qPCR para determinar el número de genomas del vector, como se describe a continuación en la sección titulada "Rendimiento alto".

- 50 Opcionalmente, el nivel de rcAAV es menor que a 1 rcAAV en 10^7 AAV recombinantes, menor que 1 rcAAV en 10^9 AAV recombinantes, o menor que 1 rcAAV en 10^{10} AAV recombinantes. Opcionalmente, el nivel de rcAAV es menor que el nivel de rcAAV producido utilizando un método equivalente, excepto que el vector plasmídico comprende dicho al menos un gen rep y al menos un gen cap. Opcionalmente, el nivel de rcAAV se mide utilizando el ensayo qPCR anterior. Opcionalmente, el nivel de rcAAV se mide usando el ensayo qPCR después de dos o tres generaciones de cultivo. Opcionalmente, el rcAAV es un rcAAV de tipo pseudosalvaje. Opcionalmente, el rcAAV es un rcAAV deficiente en cap. Opcionalmente, el rcAAV es un rcAAV deficiente en rep. Opcionalmente, el rcAAV comprende rcAAV de tipo pseudosalvaje, rcAAV deficiente en cap y rcAAV deficiente en rep.

- 60 La expresión "método equivalente al de un método de la invención" pretende referirse a un método que es idéntico, excepto que el vector plasmídico comprende dicho al menos un gen rep y al menos un gen cap.

Proporción deseada de partículas llenas a totales

- 65 También se proporciona un uso del sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico para producir una preparación de rAAV que tenga una proporción deseada de partículas llenas a totales.

También se proporciona un uso del sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico para

controlar o maximizar la proporción de partículas llenas a totales producidas durante la producción de rAAV.

También se proporciona un método para producir una preparación de rAAV que comprende:

- 5 (a) obtener el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico;
- (b) transfectar una célula huésped con el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico; y
- (c) cultivar la célula huésped en condiciones adecuadas para la producción de rAAV

en el que la preparación de rAAV comprende una proporción deseada de partículas llenas a totales.

- 10 También se proporciona un método para controlar o maximizar la proporción de partículas llenas a totales producidas durante la producción de rAAV que comprende:

- (a) obtener el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico;
- 15 (b) transfectar una célula huésped con el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico; y
- (c) cultivar la célula huésped en condiciones adecuadas para la producción de rAAV.

El método puede comprender además una etapa de recolección del rAAV para proporcionar una preparación de rAAV, que tiene opcionalmente una proporción deseada de partículas llenas a totales.

- 20 Los presentes inventores han determinado que se puede usar la alteración de la proporción de plásmido auxiliar a vector plasmídico para alterar la proporción de partículas llenas a totales que se producen. Específicamente, los inventores han determinado que al aumentar el nivel del vector plasmídico en relación con el plásmido auxiliar se reduce la proporción de partículas llenas a totales.

- 25 Las partículas "llenas" son partículas de rAAV que comprenden tanto una cápside como el genoma del vector previsto, o al menos un genoma parcial, tal como se determina usando el método de qPCR descrito a continuación. Las partículas "vacías" (es decir, las partículas que no están llenas) comprenden cápsides, pero no comprenden un genoma, o comprenden solo un genoma parcial (por lo tanto, no forman una partícula de rAAV completa) que no se detecta usando el método qPCR. Sin embargo, las preparaciones de rAAV pueden comprender tanto partículas llenas como partículas vacías. Generalmente se desea una proporción baja o minimizada de partículas vacías. Por ejemplo, si el rAAV se va a utilizar en terapia génica, cualquier partícula vacía no comprenderá el módulo de expresión de interés completo y, por lo tanto, no será eficaz en la terapia. Por otra parte, hay circunstancias en las que la presencia de partículas vacías podría ser deseable. En algunos casos, y en algunos grupos de pacientes, puede darse el caso de que las partículas vacías se comporten como "señuelos" para reducir la respuesta inmunitaria en un paciente a las partículas de rAAV administradas (documento WO2013/078400). Además, tal como se analizará con más detalle a continuación, mientras que aumentar el nivel del vector plasmídico en relación con el plásmido auxiliar disminuye la proporción de partículas llenas a totales, también aumenta el rendimiento de rAAV producido (y, por lo tanto, en algunos casos, el número total de partículas llenas que se producen). Por tanto, incluso si el usuario del método cree que es deseable una proporción alta entre partículas llenas y partículas totales, el usuario puede usar una proporción menor de plásmido auxiliar si esto conduce a un mayor rendimiento. Como alternativa, si el usuario, por ejemplo, aplica un cambio de proceso durante el desarrollo de un producto dado, por ejemplo, cambia del biorreactor A al biorreactor B, lo que da como resultado diferentes proporciones de partículas llenas a totales cuando se mantiene constante la proporción de plásmidos, la proporción de plásmidos se puede modificar para compensar o contrarrestar el efecto del cambio de proceso para garantizar una proporción constante de partículas llenas a totales durante todo el desarrollo, ya que esta proporción es un parámetro de calidad crítico de un lote de rAAV que debe controlarse y mantenerse comparable de un lote a otro.

- 50 En consecuencia, una ventaja de los métodos y usos es que permiten al usuario la flexibilidad para determinar una proporción deseable de partículas llenas a totales y modificar la proporción de plásmido auxiliar a vector plasmídico para lograr la proporción deseada de partículas llenas a totales.

- La proporción de partículas llenas a totales puede expresarse en el presente documento como el porcentaje del número total de partículas (cápsides) que teóricamente comprenden un genoma de vector o al menos un genoma parcial (suponiendo un genoma (parcial) por cápside) según se determina utilizando el siguiente ensayo de qPCR. La qPCR se lleva a cabo utilizando un par de cebadores que son capaces de amplificar al menos una región del promotor del módulo de expresión. Opcionalmente, al menos uno de los cebadores es específico (inverso y complementario o idéntico dependiendo de si el cebador es un cebador directo o un cebador inverso) para una región de al menos 12, al menos 14, al menos 16, o al menos 18 nucleótidos del promotor del módulo de expresión.
- 60 Opcionalmente, un cebador es específico para el inicio del promotor (los primeros al menos 12 nucleótidos del promotor) y el otro cebador es específico para una región del módulo de expresión que está a 150 pares de bases del sitio de unión del primer cebador.

- 65 La proporción entre partículas llenas y totales (medida como un porcentaje del número total de partículas que son partículas llenas) se puede determinar usando qPCR para determinar el número de genomas del vector (como se analizó en el párrafo anterior), y usando un ELISA específico de cápside para medir el número total de partículas.

Por ejemplo, el ELISA específico de cápside puede comprender exponer la preparación de rAAV a un anticuerpo que se une a la proteína de la cápside. Si, por ejemplo, el vector plasmídico comprende un gen cap que codifica una cápside de un serotipo AAV2, el anticuerpo puede ser un anticuerpo que se une a la cápside de AAV2. Por ejemplo, el usuario puede recubrir una placa con un anticuerpo específico para la cápside.

Luego, el usuario puede hacer pasar la preparación de rAAV sobre la superficie de la placa. Las partículas se unirán al anticuerpo y quedarán inmovilizadas en la placa. A continuación, la placa se puede lavar para eliminar los contaminantes. La cantidad de partículas presentes se puede detectar luego mediante la adición de un anticuerpo de detección que se puede unir a la cápside y está conjugado con un agente de detección, como la estreptavidina peroxidasa. La cantidad de partícula presente será proporcional al cambio de color obtenido cuando la estreptavidina peroxidasa se expone al sustrato cromogénico TMB (tetrametilbencidina).

Opcionalmente, la proporción deseada entre partículas llenas y totales (expresada como porcentaje del número total de partículas que teóricamente comprenden un genoma de vector) es de al menos un 2 %, al menos un 3 %, al menos un 4 %, al menos un 5 %, al menos un 7 %, al menos un 10 % o al menos un 15 % (medido en el producto a granel antes de la recolección, la purificación y/o la concentración de la preparación de rAAV). En muchos casos, es ventajoso aumentar la proporción de partículas llenas a totales, y los presentes inventores han determinado que el objetivo de una proporción de partículas llenas a totales de al menos un 2 %, al menos un 3 %, al menos un 4 %, al menos un 5 %, al menos un 7 %, al menos un 10 % o al menos un 15 % logra un buen equilibrio entre el mantenimiento de una alta proporción de partículas totales, además de lograr un buen rendimiento. Opcionalmente, la proporción deseada de partículas llenas a totales es una proporción de partículas llenas a totales que es al menos un 20 % o al menos un 30 % de la proporción de partículas llenas a totales lograda utilizando un método equivalente con una proporción de plásmido auxiliar a vector plasmídico de 1,8:1.

Opcionalmente, la proporción de plásmido auxiliar a vector plasmídico que se utiliza, se selecciona o se ajusta para que sea de entre 3:1 y 1:10, entre 1,5:1 y 1:9, entre 1,4:1 y 1:8, entre 1,3:1 y 1:7, entre 1,2:1 y 1:6, entre 1,1:1 y 1:5, entre 1:1 y 1:4, o entre 1:1,5 y 1:3. Opcionalmente, la proporción de plásmido auxiliar a vector plasmídico comprende un exceso molar de vector plasmídico.

Rendimiento alto

También se proporciona un uso del sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico para producir una preparación de rAAV con un rendimiento alto o deseado.

También se proporciona un uso del sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico para aumentar, optimizar o maximizar el rendimiento de rAAV producido durante la producción de rAAV.

También se proporciona un método para producir una preparación de rAAV que comprende:

- (a) obtener el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico;
- (b) transfectar una célula huésped con el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico; y
- (c) cultivar la célula huésped en condiciones adecuadas para la producción de rAAV

en el que la preparación de rAAV tiene un rendimiento alto o deseado.

También se proporciona un método para aumentar, optimizar o maximizar el rendimiento de rAAV producido durante la producción de rAAV que comprende:

- (a) obtener el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico;
- (b) transfectar una célula huésped con el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico; y
- (c) cultivar la célula huésped en condiciones adecuadas para la producción de rAAV.

El método puede comprender además una etapa de recolección del vector rAAV para proporcionar una preparación de rAAV, opcionalmente con un rendimiento alto o deseado. El término "rendimiento" se refiere a la cantidad de partículas de AAV que se preparan en los métodos o usos de la invención. El "rendimiento" puede expresarse como el número de genomas de vector (vg) por ml de medio, medido antes (medido en el producto a granel antes de la recolección, la purificación y/o la concentración de la preparación de rAAV). El número de genomas de vector se puede medir como se analiza en el encabezado "Proporción deseada de partículas llenas a totales", es decir, el rendimiento de rAAV (como las partículas de rAAV) se puede determinar usando qPCR para cuantificar el número de secuencias de ácido nucleico que comprenden un módulo que comprende una secuencia de promotor (vg).

Los presentes inventores han demostrado que se puede usar la modificación de la proporción de plásmido auxiliar a vector plasmídico para aumentar el rendimiento (es decir, la cantidad de vg o vg/ml producida) de un método que produce rAAV. Tener el mayor rendimiento posible es claramente ventajoso, ya que reduce la cantidad de recursos necesarios para producir el rAAV. Sin embargo, en algunas condiciones, aumentar el rendimiento al nivel más alto posible puede resultar en una disminución en la proporción de partículas llenas a totales. Por tanto, si el usuario está

produciendo el rAAV para una aplicación donde es importante maximizar la proporción de partículas llenas a totales, entonces puede optar por usar un rendimiento deseado que no sea el rendimiento más alto posible mediante el uso de una proporción de plásmido auxiliar a vector plasmídico que se optimiza a favor de la proporción de partículas llenas a totales (es decir, "% de llenas").

5 Tal como se utiliza en el presente documento, el término "maximizar" el rendimiento se refiere a aspirar al mayor rendimiento posible dentro de las limitaciones de los métodos de la invención. Tal como se utiliza en el presente documento, el término "optimizar" se refiere a aumentar el rendimiento, pero apuntando a un rendimiento deseado que puede no ser el rendimiento máximo posible si, por ejemplo, el usuario desea garantizar una alta proporción de
10 partículas llenas y totales (es decir, minimizar la proporción de partículas vacías).

Tal como se utiliza en el presente documento, la expresión "rendimiento alto" generalmente se refiere a un rendimiento mayor que 2×10^{10} vg/ml, o al menos 2 veces mayor que el rendimiento logrado utilizando un método equivalente con una proporción de plásmido auxiliar a vector plasmídico de 1,8:1. Un "método equivalente" es un
15 método que es idéntico, excepto que la proporción de plásmido auxiliar a vector plasmídico es de 1,8:1.

Opcionalmente, el rendimiento alto o deseado es un rendimiento mayor que 2×10^{10} vg/ml, mayor que 4×10^{10} vg/ml, mayor que 6×10^{10} vg/ml, mayor que 8×10^{10} vg/ml, mayor que 1×10^{11} vg/ml, mayor que 2×10^{11} vg/ml o mayor que 4×10^{11} vg/ml.
20

Opcionalmente, el rendimiento alto o deseado es un rendimiento que es al menos el doble, al menos 4 veces, al menos 5 veces, o al menos 6 veces mayor que el rendimiento logrado utilizando un método equivalente con una proporción de plásmido auxiliar a vector plasmídico de 1,8:1.

25 Opcionalmente, la proporción de plásmido auxiliar a vector plasmídico que se utiliza, se selecciona o se ajusta para que sea de entre 3:1 y 1:10, entre 1,5:1 y 1:9, entre 1,4:1 y 1:8, entre 1,3:1 y 1:7, entre 1,2:1 y 1:6, entre 1,1:1 y 1:5, entre 1:1 y 1:4, o entre 1:1,5 y 1:3. Opcionalmente, la proporción de plásmido auxiliar a vector plasmídico comprende un exceso molar de vector plasmídico en comparación con el plásmido auxiliar.

30 *Equilibrio entre alto rendimiento y alta proporción de partículas llenas a totales*

Como se ha analizado anteriormente, la proporción de partículas llenas a totales y el rendimiento pueden verse afectados por la alteración de la proporción de plásmido auxiliar a vector plasmídico utilizada en los métodos. Por tanto, el usuario puede ajustar la proporción de plásmido auxiliar a vector plasmídico a una proporción adecuada
35 para obtener la proporción de partículas llenas a totales deseada. Opcionalmente, los métodos o usos pueden comprender una etapa de seleccionar una proporción de plásmido auxiliar a vector plasmídico. Se puede alterar la proporción de plásmido auxiliar a vector plasmídico, o seleccionar una proporción de plásmido auxiliar a vector plasmídico, simplemente mezclando el plásmido auxiliar y el vector plasmídico en diferentes proporciones relativas.

40 El usuario puede necesitar llevar a cabo un método de prueba para determinar la proporción de plásmido auxiliar a vector plasmídico que se requiere para obtener la proporción deseada de partículas llenas a totales, pero esto no es trabajoso. Por ejemplo, los expertos en la materia pueden realizar un experimento a pequeña escala para determinar la proporción requerida de plásmido auxiliar a vector plasmídico, y luego esta proporción puede usarse en un proceso de producción a mayor escala. Opcionalmente, la etapa de seleccionar una proporción de plásmido auxiliar
45 a vector plasmídico puede comprender realizar pruebas a pequeña escala del método de la invención usando diferentes proporciones de plásmido auxiliar a vector plasmídico.

Opcionalmente, la proporción de plásmido auxiliar a vector plasmídico se selecciona o se ajusta a una proporción que logre un rendimiento equilibrado frente a la proporción de partículas llenas a totales. Opcionalmente, la proporción de plásmido auxiliar a vector plasmídico se selecciona o se ajusta a una proporción que logre un rendimiento máximo de rAAV (es decir, vg) con el rendimiento mínimo de partículas alcanzable con dicho rendimiento máximo de rAAV, es decir, el usuario maximiza el rendimiento y luego determina la proporción de plásmido auxiliar a vector plasmídico que logra ese rendimiento mientras proporciona la proporción más alta de partículas llenas a totales.
50

Opcionalmente, la proporción de plásmido auxiliar a vector plasmídico que se utiliza, se selecciona o se ajusta para que sea de entre 3:1 y 1:10, entre 1,5:1 y 1:9, entre 1,4:1 y 1:8, entre 1,3:1 y 1:7, entre 1,2:1 y 1:6, entre 1,1:1 y 1:5, entre 1:1 y 1:4, o entre 1:1,5 y 1:3. Opcionalmente, la proporción de plásmido auxiliar a vector plasmídico comprende un exceso molar de vector plasmídico.
55

60 *Transfección, cultivo y recolección*

Los métodos de la invención pueden comprender etapas de obtención del sistema de dos plásmidos de la invención, el plásmido auxiliar de la invención o el vector plasmídico, mediante la transfección de una célula huésped con el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico, y el cultivo de la célula huésped en condiciones adecuadas para la producción de AAV recombinante.
65

- La transfección de una célula huésped con el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico, puede comprender exponer la célula huésped al sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico en condiciones adecuadas para la transfección. Por ejemplo, el usuario del método puede añadir un agente de transfección (la adición de un agente de transfección se consideraría una condición adecuada para la transfección). Como alternativa, puede usarse la transfección con fosfato de calcio, la electroporación o liposomas catiónicos. Opcionalmente, la etapa de transfectar la célula huésped tiene lugar cuando la célula huésped ha crecido hasta la confluencia.
- El cultivo de la célula huésped en condiciones adecuadas para la producción de rAAV se refiere al cultivo de la célula huésped en condiciones en las que puede crecer y el AAV puede replicarse. Por ejemplo, la célula huésped se puede cultivar a una temperatura entre 32 °C y 40 °C, entre 34 °C y 38 °C, entre 35 °C y 38 °C o aproximadamente de 37 °C. Opcionalmente, la célula huésped puede cultivarse en presencia de un medio de cultivo celular completo, tal como medio de Eagle modificado por Dulbecco (DMEM). Un medio de cultivo celular completo es un medio que proporciona todos los nutrientes esenciales necesarios para el crecimiento de la célula huésped. Opcionalmente, el medio de cultivo celular completo se complementa con suero, tales como suero bovino fetal o seroalbúmina bovina.
- Opcionalmente, la célula huésped es una célula huésped como las definidas anteriormente bajo el encabezado "Célula huésped". Opcionalmente, la célula huésped se selecciona del grupo que consiste en una célula HEK293T, una célula HEK293, una célula HEK293EBNA, una célula CAP, una célula CAP-T, una célula AGE1.CR, una célula PerC6, una célula C139 y una célula EB66. Opcionalmente, la célula huésped es una célula que expresa una proteína E1A/B funcional.
- El método puede comprender además una etapa de purificar el rAAV. En general, una etapa de purificación del rAAV implicará aumentar la concentración del rAAV en comparación con otros componentes de la preparación. Opcionalmente, la etapa de purificación del rAAV da como resultado una preparación concentrada de rAAV. Opcionalmente, la etapa de purificación del rAAV da como resultado un rAAV aislado.
- Puede usarse cualquier método de purificación adecuado. Opcionalmente, la etapa de purificación del rAAV se lleva a cabo utilizando una técnica seleccionada del grupo que consiste en centrifugación en gradiente de densidad (como la centrifugación en gradiente de densidad con CsCl o iodixanol), filtración, cromatografía de intercambio iónico, cromatografía de exclusión por tamaño molecular, cromatografía de afinidad y cromatografía de interacción hidrofóbica.
- Opcionalmente, el método comprende concentrar aún más el rAAV usando ultracentrifugación, filtración de flujo tangencial o filtración en gel.
- Opcionalmente, el método comprende formular el rAAV con un excipiente farmacéuticamente aceptable. Los excipientes farmacéuticamente aceptables pueden comprender vehículos, diluyentes y/u otros agentes medicinales, agentes farmacéuticos o adyuvantes, etc. Opcionalmente, los excipientes farmacéuticamente aceptables comprenden solución salina. Opcionalmente, los excipientes farmacéuticamente aceptables comprenden seroalbúmina humana.
- Preparaciones de AAV recombinante (rAAV)**
- Puede obtenerse una preparación de AAV recombinante (rAAV) mediante un método de la invención. La preparación de AAV recombinante se puede obtener mediante un método de la invención.
- Las preparaciones de AAV recombinantes obtenibles u obtenidas por los métodos de la invención son ventajosas, ya que generalmente comprenden un nivel bajo de rcAAV y/o tienen una proporción deseada de partículas llenas a totales. Las secciones tituladas "Proporción deseada de partículas llenas a totales", "Equilibrio entre alto rendimiento y alta proporción de partículas llenas a totales" y "Nivel bajo de AAV competente para la replicación (rcAAV)" proporcionan más detalles de lo que significan las expresiones nivel bajo de rcAAV y una proporción deseada de partículas llenas a totales.

Tabla de listado de secuencias

Número de identificación de la secuencia	Secuencia
1	Genoma de AAV2 (AF043 3 03.1)
2	Genoma de AdV5 (AC_000008.1)
3	Genoma de AdV2 (AC_000007.1)
4	Región de gen auxiliar
5	Genoma alternativo de AdV5 (AY339865.1)

Ejemplos

Ejemplo 1: Construcción de plásmidos auxiliares y vectores de un sistema de dos plásmidos ("transdivididos")

5 *Plásmido auxiliar*

Para preparar el plásmido auxiliar, los nucleótidos 200-4497 de AAV2 de tipo salvaje (número de acceso de Genbank AF043303; SEQ ID NO:1) que contenían los genes rep y cap se clonaron en el pUC19 (Yanisch-Perron *et al.* (1985), Gene, 33:103-119). A continuación, se delecionaron los nucleótidos 4461-4497 de AAV2 para minimizar la

10 homología de secuencia con el vector plasmídico (cuya construcción se describe a continuación). Para evitar la expresión de Rep 78 mientras se mantiene la expresión de Rep 68, se eliminó el intrón dentro del gen rep clonado. Para proporcionar la expresión de Rep 52 después de la eliminación del intrón, los nucleótidos de AAV2 correspondientes a Rep 52, incluido el promotor p19, se clonaron inmediatamente 3' del gen Rep 68 sin intrones. A continuación, se delecionaron la mayoría de las secuencias del gen cap.

15 Los dos promotores p40 (uno en cada una de las duplicaciones del gen rep que codifica Rep 68 y Rep 52) se volvieron no funcionales mediante la ablación de las cajas TATA (mutación de la T correspondiente a la posición 1823 de AAV2, y AAG correspondiente a las posiciones de AAV2 1826-1828 a C y CTC, respectivamente).

20 El "módulo de rep" resultante se clonó luego en un plásmido que comprende un tramo de 8342 nucleótidos que comprende los genes VA ARN I y II, E2A y E4 funcionales de adenovirus (es decir, virus auxiliar) serotipo 5 (SEQ ID NO:4). El esqueleto del plásmido, que contiene un gen de resistencia a la kanamicina y un origen de replicación bacteriano, tenía aproximadamente 2,2 kb de longitud, dando como resultado un plásmido auxiliar de 14021 nucleótidos.

25 La figura 2 es un esquema del plásmido auxiliar que muestra las características principales. Para el "módulo de rep", se indican las partes de la secuencia AAV2, por referencia a las posiciones de nucleótidos de SEQ ID NO:1.

30 *Vector plasmídico*

Para preparar el vector plasmídico, los nucleótidos 200-4497 de AAV2 de tipo salvaje (número de acceso de Genbank AF043303; SEQ ID NO:1) que contenían los genes rep y cap se clonaron en el pUC19. Dos porciones de la secuencia del gen rep, entre los promotores p5 y p19 y entre los promotores p19 y p40, respectivamente, se eliminaron después para evitar la expresión de la proteína Rep mientras se mantenía el gen cap secuencia abajo

35 bajo la regulación de los tres promotores nativos. En cuanto al plásmido auxiliar descrito anteriormente, Los nucleótidos 4461-4497 de AAV2 se delecionaron para minimizar la homología de secuencia con el plásmido auxiliar.

Para minimizar o prevenir la traducción de productos no deseados de posibles codones de iniciación dentro de la región promotora restante, se eliminaron cuatro codones ATG y un codón GTG: los ATG en las posiciones correspondientes a los nucleótidos 321-323, 955-957 y 993-995 de AAV2, y un GTG correspondiente a los

40 nucleótidos 1014-1016 de AAV2, fueron delecionados, mientras que ATG en los nucleótidos 766-768 de AAV2 se mutó a ATT.

El gen cap de AAV2 que codifica las VP 1, 2 y 3 inmediatamente 3' de la región promotora anterior se reemplazó por la secuencia correspondiente de un gen cap manipulado, y este gen cap es 3 nucleótidos (equivalente a un aminoácido codificado adicional) más largo que el gen cap de AAV2. El módulo de "promotor-cap" resultante se clonó en un esqueleto del plásmido que contenía un gen de resistencia a la kanamicina y un origen bacteriano de replicación. En este esqueleto se insertó un módulo de expresión, que contenía la secuencia transgénica unida al promotor y los elementos reguladores de la transcripción de poliA, flanqueada por la secuencia de AAV2 que comprende las ITR nativas de AAV2 (nucleótidos 1-145 y 4535-4679 de AAV2).

50

En el caso del vector plasmídico que comprende el módulo de expresión del factor IX de 2672 nucleótidos (ITR-a-ITR), como se usa en el ejemplo 2, el esqueleto del plásmido tenía una longitud aproximada de 2,3 kb, dando como resultado un vector plasmídico de 8525 nucleótidos. En todas las actividades posteriores, incluida la producción de los lotes de rAAV del ejemplo 3, el esqueleto del plásmido se redujo en longitud a aproximadamente 2 kb. El módulo de expresión de α -galactosidasa A (GLA) (ejemplo 3) tenía una longitud de 2297 nucleótidos.

55

La figura 3 es un esquema del vector plasmídico que muestra las características principales. Para el módulo de "promotor-cap", se indican las porciones de la secuencia AAV2, por referencia a las posiciones de nucleótidos de SEQ ID NO:1.

60

Ejemplo 2: Sistema de dos plásmidos: comparación de proporciones de plásmido auxiliar:vector

Cultivo de células

65 Las células HEK293T se mantuvieron en cultivo adherente en condiciones convencionales a 37 °C, 95 % de

humedad relativa y 5 % v/v de CO₂ en medio de Eagle modificado por Dulbecco (DMEM) complementado con suero fetal bovino (FBS) al 10 % y GlutaMax™ (dipéptido de L-alanina-L-glutamina) al 1 %. La confluencia celular durante la transferencia varió entre el 40 y el 95 %.

5 Transfección de células HEK293T y preparación de lisados celulares

Las células HEK293T se transfectaron con un plásmido auxiliar y un vector plasmídico (este último comprende el gen cap manipulado y el módulo de expresión del factor IX; ejemplo 1) usando diferentes proporciones molares de plásmido (de auxiliar:vector 3:1 a 1:6) mientras se mantiene la cantidad total de ADN plasmídico. Se sembraron 1,5 × 10⁵ células viables por cm² área de cultivo en placas de 6 cm en un volumen de 3 ml de DMEM, FBS al 10 %, GlutaMax™ al 1 % el día anterior a la transfección, lo que da como resultado una confluencia del 60-70 % el día de la transfección. Los complejos de PEI-ADN se prepararon en DMEM sin complementos utilizando el reactivo de transfección de polietilimina lineal PEIpro™ (Polyplus) según el manual del fabricante. Se mantuvo una cantidad de 6 µg de ADN plasmídico total y una proporción de PEI a ADN de 2:1 independientemente de las proporciones de combinación de plásmidos aplicadas. Las células se cultivaron hasta el día 3 después de la transfección, se recolectaron del medio y se lisaron mediante tres ciclos de congelación-descongelación (-80 °C y 37 °C). Los residuos celulares se retiraron por centrifugación a 10 000 x g durante 5 minutos.

20 Cuantificación de los genomas de vectores rAAV

El ensayo del genoma del vector AAV se basa en una reacción en cadena de la polimerasa cuantitativa (qPCR) específica para la secuencia de promotor del módulo de expresión de rAAV. En principio, los cebadores de qPCR se pueden diseñar para unirse a cualquier parte del genoma de AAV recombinante que no sea común a los genomas de AAV de tipo salvaje, pero se recomienda no usar secuencias de plantilla de cebadores muy cercanas a las ITR, ya que hacerlo puede conducir a una medición exagerada del título del genoma del vector.

Las muestras de ensayo de lisado celular se sometieron a un procedimiento de tratamiento con nucleasas para eliminar los genomas del vector no encapsidados antes de realizar la qPCR. Para ello, las muestras se prediluyeron 1:250 en agua sin nucleasas que contenía Pluronic F-68 al 0,125 %. Se utilizaron 25 µl de la predilución para la digestión con 2 unidades de Turbo DNase (ThermoFisher Scientific, Waltham, EE. UU.) y 1x tampón de reacción de Turbo DNase, dando como resultado un volumen de reacción total de 29 µl. La incubación se realizó durante 1 h a 37 °C. Después, se añadió 1 volumen de NaOH 0,4 M y las muestras se incubaron durante 45 min a 65 °C. Se añadieron 1035 µl de agua sin nucleasas complementada con Pluronic F-68 al 0,1 % junto con 30 µl de HCl 0,4 M. Para controlar la calidad del digerido de Turbo DNase, se midieron en paralelo un control de tendencia que contenía lisado celular no purificado con un título conocido de genoma de vector AAV y controles añadidos que usaban ADN plasmídico que portaba la secuencia de promotor.

En cada muestra, se mezclaron 12,5 µl de QuantiFast SYBR Green PCR Master Mix (Qiagen, Venlo, Países Bajos) con 0,75 µl de solución madre final de cebador de qPCR (que contenía 10 µM de cada cebador) y se llevaron hasta un volumen de 20 µl con agua sin nucleasas. Se añadieron a la mezcla 5 µl de lisado celular tratado con Turbo DNase o muestra de ensayo de virus purificado (volumen total de reacción 25 µl, concentración de cebador final en la reacción 300 nM cada uno) y se realizó una qPCR en un ciclador de PCR en tiempo real CFX 96 Touch (Bio-Rad Laboratories Inc., Hercules, EE. UU.) con las siguientes etapas del programa: 95 °C 5 min; 39 ciclos (95 °C 10 s, 60 °C 30 s, lectura de placa); 95 °C 10 s; 60-95 °C (+0,5 °C/etapa), 10 s; lectura de placa. Para controlar la calidad de la qPCR, se midió en paralelo un control de tendencia con un título conocido del genoma del vector AAV. Para comprobar si hay contaminaciones, también se incluyó un control sin plantilla (NTC, 5 µl de H₂O). La fila del patrón, las muestras de ensayo y los controles se midieron por triplicado para cada dilución. Las muestras de ensayo de virus purificadas y el control de tendencia se midieron generalmente a 3 diluciones diferentes en tampón EB (Tris-Cl 10 mM, pH 8,5). Las muestras de ensayo de lisado celular tratado con Turbo DNase se usaron directamente en la qPCR sin ninguna dilución adicional. Los datos se analizaron utilizando el software CFX Manager™ 3.1 (Bio-Rad Laboratories Inc.).

El análisis de la curva de fusión confirmó la presencia de un solo amplicón. La amplificación da como resultado amplicones de ADN bicatenario nacientes detectados con el intercalador fluorescente SYBR Green para controlar la reacción de PCR en tiempo real. Cantidades conocidas del material genético del promotor, en forma de plásmido linealizado, se diluyeron en serie para crear una curva patrón, y el título del genoma del vector de muestra se interpoló a partir de la curva patrón.

60 Cuantificación de partículas de rAAV (cápsides)

El método ELISA de titulación de AAV2 es una medida del total de partículas de AAV (cápsides) y se basa en un kit disponible en el mercado (Progen™, Heidelberg, Alemania; número de catálogo PRATV). Esta técnica inmunométrica de tipo sándwich utiliza el anticuerpo monoclonal A20 (Wobus *et al.* (2000), J. Virol., 74:9281-9293) tanto para la captura como para la detección. El anticuerpo es específico para un epítipo conformacional presente en las cápsides ensambladas de los serotipos AAV2, AAV3 y la cápside manipulada utilizada en estos experimentos.

El kit ELISA de titulación de AAV2 se utilizó para cuantificar las partículas de AAV totales en lisados celulares y preparaciones de virus purificados según las instrucciones del fabricante. En resumen, se añadieron 100 µl de control de kit AAV2 diluido, muestras de ensayo o control de tendencias de la cápside manipulada con un título conocido de partículas totales por pocillo de una placa de microtitulación recubierta con anticuerpo monoclonal A20 y se incubaron durante 1 h a 37 °C. La fila del patrón, las muestras de ensayo y los controles se midieron por duplicado para cada dilución. En una segunda etapa, se añadieron 100 µl de anticuerpo monoclonal A20 conjugado con biotina prediluido (1:20 en tampón de ensayo [ASSB]) y se incubó durante 1 hora a 37 °C. Después, se añadieron 100 µl de un conjugado de estreptavidina peroxidasa prediluido (1:20 en ASSB) y se incubó durante 1 hora a 37 °C. Se añadieron 100 µl de solución de sustrato (TMB [tetrametilbencidina]) y, después de incubar durante 15 min, la reacción se detuvo usando 100 µl de solución de parada. La absorbancia se midió fotoquímicamente a 450 nm utilizando el lector de microplacas SpectraMax M3 (Molecular Devices, San José, EE.UU.). Los datos se analizaron con el software SoftMax Pro 7.0 (Molecular Devices).

Las muestras de ensayo se diluyeron en el Intervalo de ensayo y las concentraciones de partículas totales de AAV se determinaron por interpolación usando la curva patrón que se preparó usando el control del kit AAV2 provisto. Se usó ASSB como blanco.

Proporción de genoma de vector a partículas totales

La proporción entre los genomas de vector y el total de partículas de AAV se expresa como un porcentaje. Esto se basa en el título del genoma del vector (determinado por qPCR, como se describió anteriormente) y el número total de partículas de AAV (determinado por ELISA de cápside, como se describió anteriormente).

Resultados

Como se desprende de la figura 5A y B, el aumento de la proporción de vector plasmídico condujo a mayores rendimientos tanto de partículas como de genoma del vector. Sin embargo, el aumento relativo en los rendimientos de partículas (cápside) fue más pronunciado y la meseta en los rendimientos del genoma del vector se alcanzó antes. Estas observaciones fueron responsables de una disminución gradual en la proporción entre el genoma del vector y las partículas totales (figura 5C). Una proporción de plásmido auxiliar:vector 1:3 dio como resultado un aumento casi máximo en el título del genoma del vector de aproximadamente 3,5 veces en comparación con la proporción de 1,8:1 aplicada anteriormente, mientras que el título de partículas aumentó aproximadamente 8 veces.

Con fines comparativos, dos plásmidos en una configuración convencional "no dividida" (es decir, en la que un plásmido contenía las mismas secuencias del genoma auxiliar y del vector de AdV (módulo de expresión del factor IX), y el otro plásmido contenía la misma secuencia del gen cap, además de un módulo rep de AAV2 que contenía los cuatro genes rep (de modo que las funciones Rep y Cap no se dividen entre los dos plásmidos) se transfectaron en una proporción de plásmido molar establecida de 1,6:1, revelando rendimientos del genoma del vector y de partículas considerablemente más bajos.

La secuenciación confirmatoria posterior del plásmido auxiliar reveló una mutación en el codón dentro de la secuencia que codifica Rep 68 correspondiente a las posiciones 429-431 de AAV2 (SEQ ID NO:1), lo que produjo una sustitución de leucina por fenilalanina en la posición del aminoácido 37 del proteína Rep 68. La corrección de la secuencia del plásmido auxiliar para restaurar el codón AAV2 de tipo salvaje que codifica leucina dio como resultado un mayor rendimiento de vg/ml y un aumento de aproximadamente el doble en la proporción entre los genomas del vector y las partículas totales (es decir, % de partículas llenas), cuando se compara directamente con el plásmido auxiliar "mutado". La secuencia corregida del plásmido auxiliar se usó en todos los trabajos posteriores, incluyendo la producción de los lotes de rAAV analizados en el ejemplo 3.

Ejemplo 3: Determinación de la frecuencia de AAV competente para la replicación (rcAAV) en rAAV producido por un sistema de dos plásmidos transdivididos

Pruebas de rcAAV

Se sometió la sustancia farmacológica a granel de rAAV purificada, fabricada a gran escala utilizando el biorreactor iCellis® con células transfectadas con el sistema de dos plásmidos y purificadas mediante una serie de procesos posteriores para eliminar las impurezas del producto, a una prueba del límite para la presencia de rcAAV. Los lotes de rAAV se produjeron usando un vector plasmídico que comprende el gen cap manipulado y (i) un módulo de expresión de α-galactosidasa A (como se menciona en el ejemplo 1) o (ii) un módulo de expresión del factor IX idéntico al mencionado en el ejemplo 1, excepto por una secuencia codificante del factor IX diferente con codones optimizados parcialmente. El sistema de dos plásmidos utilizado para producir estos lotes de rAAV utilizó el esqueleto de plásmido de vector acortado y la secuencia de plásmido auxiliar corregida como se menciona en los ejemplos 1 y 2, respectivamente. En la prueba del límite, las células HEK293 son transducidas por el rAAV en su forma más concentrada (después de la purificación y de la preformulación del fármaco) en presencia o ausencia de adenovirus de tipo salvaje. Se llevan a cabo tres rondas sucesivas de amplificación del virus como se describe a continuación.

Las células se siembran en matraces T75/T175. A un 80 % +/- 10 de confluencia y >90 % de viabilidad, las células se transducen con y sin adenovirus auxiliar de tipo salvaje serotipo 5 ("Ad5wt") a una multiplicidad de infección ("multiplicity of infection", MOI) de $7,8 \times 10^3$ y el producto rAAV a analizar. Después de 2 días de incubación, las células se recolectan separándolas mecánicamente. Se recolectan y centrifugan tres conjuntos de células de la ronda uno: 1) uno se guarda de reserva (se almacena en caso de que se requiera repetir el análisis); 2) uno para la purificación y el análisis del ADN; 3) una para una segunda ronda de transducción. Las células para la purificación y las siguientes rondas de transducción se lisan mediante congelación/descongelación (3x) en hielo seco/etanol al baño maría a 37 °C, y Ad5wt se inactiva mediante incubación a 56 °C durante 30 minutos y centrifugación a 1500 g durante 2 minutos. Las células sembradas para la segunda transducción se transducen con lisado de la primera cosecha, con y sin Ad5wt, y se repite todo el proceso para recolectar muestras de la cosecha 2. Las células sembradas para la tercera transducción se transducen con lisado de la segunda cosecha, con y sin Ad5wt y se repite todo el proceso para recolectar muestras de la cosecha 3.

El ADN genómico se extrae de cada una de las tres etapas de amplificación utilizando el kit de tejido DNeasy® (Qiagen) y se cuantifica mediante UV a A260/A280. La presencia de rcAAV se detecta mediante PCR cuantitativa en tiempo real (qPCR): El ADN se aísla de las células HEK293 y se amplifican dos secuencias a partir del ADN aislado utilizando 1) un cebador específico para la secuencia codificante de rep de AAV2 ("rep"), y 2) un gen constitutivo endógeno en las células HEK293 (albúmina humana; hAlb).

Se añade ADN a una concentración de 100 ng a la reacción de qPCR de 25 µl. El número de copias de rep se calcula en relación con el intervalo patrón del plásmido (de 100 a 1×10^8 copias por reacción de qPCR) y se calcula el número relativo de copias de la secuencia rep por célula como la proporción de las copias de rep y las copias de la secuencia de albúmina humana, multiplicado por dos (hAlb, dos copias/célula). El aumento del nivel de la secuencia rep por célula en rondas de amplificación consecutivas indican la presencia de rcAAV.

El control positivo es AAV2 de tipo salvaje con adenovirus (Ad5wt); se ensayan solos o en presencia de la muestra del producto de rAAV como controles de inhibición. El control confirma que el límite de detección para la prueba (LOD) está entre 10 rcAAV 1×10^{10} y 10 rcAAV por 1×10^{11} vg de muestra de producto rAAV para AAV2. El LOD es un parámetro dependiente del ensayo y del producto. Los controles negativos consisten en células no infectadas (con o sin adenovirus) y células transducidas con AAV2 de tipo salvaje (sin adenovirus). La replicación se establece cuando el número de copias de la secuencia rep por célula es >10 en al menos 1 de las 3 rondas de amplificación, y en rondas posteriores si hay más de una. La ausencia de rep se indica como "sin replicación", lo que significa <10 rcAAV en 1×10^{10} a 1×10^{11} vg de muestra de ensayo. La detección de la secuencia rep se indica como "replicación", es decir, la detección de rcAAV en 1×10^{10} a 1×10^{11} vg de muestra de ensayo.

Resultados

La prueba del límite se realizó en varios lotes de cada uno de los productos rAAV que contenían α-galactosidasa A y factor IX. Los controles positivos competentes para la replicación de la coinfección de AAV2 de tipo salvaje con Ad5wt confirmaron la detección de rcAAV por el ensayo. Los controles positivos de rcAAV (AAV2) añadidos a la muestra de producto de rAAV no demostraron inhibición de la muestra. Los controles positivos confirmaron que el límite de detección para la prueba (LOD) fue de 10 rcAAV por 1×10^{11} vg de muestra de producto de rAAV que contiene factor IX, y 10 rcAAV por 1×10^{10} o 1×10^{11} vg de muestra de producto de rAAV que contiene α-galactosidasa A. Los controles negativos de células no infectadas (con o sin adenovirus) y las células transducidas con AAV2 de tipo salvaje (sin adenovirus) fueron negativos para rcAAV.

Por tanto, se demostró que los lotes de productos de rAAV ensayados producidos con el sistema de dos plásmidos contenían <1 rcAAV por 1×10^{10} vg de producto de rAAV que contiene factor IX, y <1 rcAAV por 1×10^9 - 1×10^{10} vg del producto de rAAV que contiene α-galactosidasa A.

Ejemplo 4

El sistema de dos plásmidos se usó en relación con otros transgenes y cápsides, y los datos pertinentes se presentan en las figuras 6-9.

LISTADO DE SECUENCIAS

<110> FREELINE THERAPEUTICS LIMITED

<120> SISTEMA DE PLASMIDO

<130> N415537EP

<160> 5

<170> PatentIn versión 3.5

<210> 1

<211> 4679

<212> ADN

<213> Virus adenoasociado de tipo 2

<400> 1

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc	60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg	120
gccaactcca tcactagggg ttcctggagg ggtggagtcg tgacgtgaat tacgtcatag	180
ggttagggag gtcctgtatt agaggtcacg tgagtgtttt gcgacatttt gcgacaccat	240
gtggtcacgc tgggtattta agcccgagtg agcacgcagg gtctccattt tgaagcggga	300
ggtttgaacg cgcagccgcc atgccggggt tttacgagat tgtgattaag gtccccagcg	360
accttgacga gcatctgccc ggcattttctg acagctttgt gaactgggtg gccgagaagg	420
aatgggagtt gccgccagat tctgacatgg atctgaatct gattgagcag gcacccctga	480
ccgtggccga gaagctgcag cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc	540
cggaggccct tttctttgtg caatttgaga agggagagag ctacttccac atgcacgtgc	600
tcgtggaaac caccggggtg aaatccatgg ttttgggacg tttcctgagt cagattcgcg	660
aaaaactgat tcagagaatt taccgcggga tcgagccgac tttgccaaac tggttcgcg	720
tcacaaagac cagaaatggc gccggaggcg ggaacaagg tgggatgag tgctacatcc	780
ccaattactt gctccccaaa acccagcctg agctccagtg ggcgtggact aatatggaac	840
agtatttaag cgcctgtttg aatctcacgg agcgtaaagc gttggtggcg cagcatctga	900
cgcacgtgtc gcagacgcag gagcagaaca aagagaatca gaatcccaat tctgatgcgc	960
cgggtgatcag atcaaaaact tcagccaggt acatggagct ggtcgggtgg ctctgggaca	1020
aggggattac ctcggaagaag cagtggatcc aggaggacca ggcctcatal atctccttca	1080
atgcggcctc caactcgcg tcccaaatca aggctgcctt ggacaatgcg ggaaagatta	1140
tgagcctgac taaaaccgcc ccgactacc tgggtgggcca gcagcccgtg gaggacattt	1200
ccagcaatcg gatattataa attttggaa taaacgggta cgatcccaa tatgcggctt	1260
ccgtctttct gggatgggcc acgaaaaagt tcggcaagag gaacaccatc tggctgtttg	1320
ggcctgcaac taccgggaag accaacatcg cggaggccat agcccacact gtgcccttct	1380

acgggtgcgt	aaactggacc	aatgagaact	ttcccttcaa	cgactgtgtc	gacaagatgg	1440
tgatctggtg	ggaggagggg	aagatgaccg	ccaaggtcgt	ggagtcggcc	aaagccattc	1500
tcggaggaag	caaggtgcgc	gtggaccaga	aatgcaagtc	ctcggccag	atagaccga	1560
ctcccgat	cgtcacctcc	aacaccaaca	tgtgcgcgt	gattgacggg	aactcaacga	1620
ccttcgaaca	ccagcagccg	ttgcaagacc	ggatgttcaa	atttgaactc	acccgccgtc	1680
tggatcatga	ctttgggaag	gtcaccaagc	aggaagtcaa	agactttttc	cggtgggcaa	1740
aggatcacgt	ggttgaggtg	gagcatgaat	tctacgtcaa	aaaggggtga	gccaaagaaa	1800
gacccgcccc	cagtgcgcga	gatataagtg	agcccaaacg	ggtgcgcgag	tcagttgcgc	1860
agccatcgac	gtcagacgcg	gaagcttcga	tcaactacgc	agacaggtac	caaaacaaat	1920
gttctcgtca	cgtgggcatg	aatctgatgc	tgtttccctg	cagacaatgc	gagagaatga	1980
atcagaattc	aaatatctgc	ttcactcagc	gacagaaaga	ctgttttagag	tgctttcccg	2040
tgtcagaatc	tcaaccgcgt	tctgtcgtca	aaaaggcgta	tcagaaactg	tgctacattc	2100
atcatatcat	gggaaaggtg	ccagacgctt	gcactgcctg	cgatctggtc	aatgtggatt	2160
tggatgactg	catctttgaa	caataaatga	tttaaatacag	gtatggctgc	cgatggttat	2220
cttcagatt	ggctcgagga	cactctctct	gaaggaataa	gacagtgggtg	gaagctcaaa	2280
cctggccccc	caccacccaa	gcccgcagag	cggcataagg	acgacagcag	gggtcttgtg	2340
cttcctgggt	acaagtacct	cggacccttc	aacggactcg	acaagggaga	gccggtcaac	2400
gaggcagacg	ccgcggccct	cgagcacgac	aaagcctacg	accggcagct	cgacagcgga	2460
gacaacccgt	acctcaagta	caaccacgcc	gacgcggagt	ttcaggagcg	ccttaaagaa	2520
gatacgtctt	ttggggggcaa	cctcggacga	gcagtcttcc	aggcgaaaaa	gagggttctt	2580
gaacctctgg	gcctgggttg	ggaacctgtt	aagacggctc	cgggaaaaaa	gaggccggta	2640
gagcactctc	ctgtggagcc	agactcctcc	tcgggaaccg	gaaagggcgg	ccagcagcct	2700
gcaagaaaaa	gattgaattt	tggtcagact	ggagacgcag	actcagtacc	tgacccccag	2760
cctctcggac	agccaccagc	agccccctct	ggtctgggaa	ctaatacgat	ggctacaggc	2820
agtggcgcac	caatggcaga	caataacgag	ggcgccgacg	gagtgggtaa	ttcctcggga	2880
aattggcatt	gcgattccac	atggatgggc	gacagagtca	tcaccaccag	caccgaacc	2940
tgggccctgc	ccacctacaa	caaccacctc	tacaaacaaa	tttcagcca	atcaggagcc	3000
tcgaacgaca	atcactactt	tggttacagc	accccttggg	ggtattttga	cttcaacaga	3060
ttccactgcc	acttttcacc	acgtgactgg	caaagactca	tcaacaacaa	ctggggattc	3120
cgacccaaga	gactcaactt	caagctcttt	aacattcaag	tcaaagaggt	cacgcagaat	3180
gacggtacga	cgacgattgc	caataacctt	accagcacgg	ttcaggtgtt	tactgactcg	3240

gagtaccagc tcccgtacgt cctcggctcg gcgcatcaag gatgcctccc gccgttccca 3300
gcagacgtct tcatggtgcc acagtatgga tacctcacc tgaacaacgg gagtcaggca 3360
gtaggacgct cttcatttta ctgcctggag tactttcctt ctcagatgct gcgtaccgga 3420
aacaacttta ccttcagcta cacttttgag gacgttcctt tccacagcag ctacgctcac 3480
agccagagtc tggaccgtct catgaatcct ctcacgacc agtacctgta ttacttgagc 3540
agaacaaaca ctccaagtgg aaccaccacg cagtcaaggc ttcagttttc tcaggccgga 3600
gcgagtgaca ttccgggacca gtctaggaac tggcttcctg gaccctgtta ccgccagcag 3660
cgagtatcaa agacatctgc ggataacaac aacagtgaat actcgtggac tggagctacc 3720
aagtaccacc tcaatggcag agactctctg gtgaatccgg gcccggccat ggcaagccac 3780
aaggacgatg aagaaaagtt ttttcctcag agcgggggtt tcatctttgg gaagcaaggc 3840
tcagagaaaa caaatgtgga cattgaaaag gtcattgatta cagacgaaga ggaaatcagg 3900
acaaccaatc ccgtggctac ggagcagtat ggttctgtat ctaccaacct ccagagaggc 3960
aacagacaag cagctaccgc agatgtcaac acacaaggcg ttcttccagg catggtctgg 4020
caggacagag atgtgtacct tcaggggccc atctgggcaa agattccaca cacggacgga 4080
cattttcacc cctctccct catgggtgga ttccggactta aacaccctcc tccacagatt 4140
ctcatcaaga acaccccggt acctgcgaat ccttcgacca ccttcagtgc ggcaaagttt 4200
gcttccttca tcacacagta ctccacggga caggtcagcg tggagatcga gtgggagctg 4260
cagaaggaaa acagcaaacg ctggaatccc gaaattcagt aactttcaa ctacaacaag 4320
tctgttaatg tggactttac tgtggacact aatggcgtgt attcagagcc tcgccccatt 4380
ggcaccagat acctgactcg taatctgtaa ttgcttgta atcaataaac cgtttaattc 4440
gtttcagttg aactttggtc tctgcgtatt tctttcttat ctagtttcca tggctacgta 4500
gataagtagc atggcggggt aatcattaac tacaaggaac ccctagtgat ggagttggcc 4560
actccctctc tgcgcgctcg ctcgctcact gaggcggggc gaccaaaggt cggccgacgc 4620
ccgggctttg cccggggcgg ctcagtgagc gagcgagcgc gcagagaggg agtggccaa 4679

<210> 2

<211> 35938

<212> ADN

<213> Adenovirus de tipo 5

<400> 2

catcatcaat aatatacctt attttggtt gaagccaata tgataatgag ggggtggagt 60
ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtgtg gcggaagtgt 120
gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180
gtgtgcgcg gtgtacacag gaagtgacaa ttttcgcgcg gttttaggcg gatgttgtag 240

taaatttggg	cgtaaccgag	taagatttgg	ccattttcgc	gggaaaactg	aataagagga	300
agtgaaatct	gaataatttt	gtgttactca	tagcgcgtaa	tatttgtcta	gggccgcggg	360
gactttgacc	gtttacgtgg	agactcgccc	aggtgttttt	ctcaggtggt	ttccgcgttc	420
cgggtcaaag	ttggcgtttt	attattatag	tcagctgacg	tgtagtgtat	ttatacccg	480
tgagttcctc	aagaggccac	tcttgagtgc	cagcgagtag	agttttctcc	tccgagccgc	540
tccgacaccg	ggactgaaaa	tgagacatat	tatctgccac	ggaggtgtta	ttaccgaaga	600
aatggccgcc	agtcttttgg	accagctgat	cgaagaggta	ctggctgata	atcttccacc	660
tcctagccat	tttgaaccac	ctacccttca	cgaactgtat	gatttagacg	tgacggcccc	720
cgaagatccc	aacgaggagg	cggtttcgca	gatttttccc	gactctgtaa	tgttggcggg	780
gcaggaaggg	attgacttac	tcacttttcc	gccggcgccc	ggttctccgg	agccgcctca	840
cctttcccgg	cagcccgagc	agccggagca	gagagccttg	ggtccgggtt	ctatgccaaa	900
ccttgtagcg	gaggtgatcg	atcttacctg	ccacgaggct	ggctttccac	ccagtgcga	960
cgaggatgaa	gagggtgagg	agtttgtggt	agattatgtg	gagcaccctg	ggcacgggtg	1020
caggtcttgt	cattatcacc	ggaggaatac	gggggaccca	gatattatgt	gttcgctttg	1080
ctatatgagg	acctgtggca	tgtttgtcta	cagtaagtga	aaattatggg	cagtgggtga	1140
tagagtgggtg	ggtttgggtg	ggtaattttt	tttttaattt	ttacagtttt	gtggtttaaa	1200
gaattttgta	ttgtgatttt	tttaaaaggt	cctgtgtctg	aacctgagcc	tgagcccgag	1260
ccagaaccgg	agcctgcaag	acctaccgcg	cgtcctaata	tggcgccctg	tatcctgaga	1320
cgcccgcacat	cacctgtgtc	tagagaatgc	aatagtagta	cggatagctg	tgactccggt	1380
ccttctaaca	cacctcctga	gatacacccg	gtgggtcccg	tgtgccccat	taaaccagtt	1440
gccgtgagag	ttgggtgggcg	tcgccaggct	gtggaatgta	tcgaggactt	gcttaacgag	1500
cctgggcaac	ctttggactt	gagctgtaaa	cgccccaggc	cataagggtg	aaacctgtga	1560
ttgcgtgtgt	ggttaacgcc	tttgtttgct	gaatgagttg	atgtaagttt	aataaagggg	1620
gagataatgt	ttaacttgca	tggcgtgtta	aatggggcgg	ggcttaaagg	gtatataatg	1680
cgccgtgggc	taatcttggg	tacatctgac	ctcatggagg	cttgggagtg	tttggagat	1740
ttttctgctg	tgcgtaactt	gctggaacag	agctctaaca	gtacctcttg	gttttgagg	1800
tttctgtggg	gctcatccca	ggcaaagtta	gtctgcagaa	ttaaggagga	ttacaagtgg	1860
gaatttgaag	agcttttgaa	atcctgtggg	gagctgtttg	attctttgaa	tctgggtcac	1920
caggcgcttt	tccaagagaa	ggtcatcaag	actttggatt	tttccacacc	ggggcgcgct	1980
gcggctgctg	ttgctttttt	gagttttata	aaggataaat	ggagcgaaga	aaccatctg	2040
agcggggggg	acctgctgga	ttttctggcc	atgcatctgt	ggagagcggg	tgtgagacac	2100
aagaatcgcc	tgctactgtt	gtcttcogtc	cgccccgcga	taataccgac	ggaggagcag	2160

cagcagcagc	aggaggaagc	caggcggcgg	cggcaggagc	agagcccatg	gaacccgaga	2220
gccggcctgg	accctcgga	atgaatgttg	tacaggtggc	tgaactgtat	ccagaactga	2280
gacgcatttt	gacaattaca	gaggatgggc	aggggctaaa	gggggtaaag	agggagcggg	2340
gggcttgtga	ggctacagag	gaggctagga	atctagcttt	tagcttaatg	accagacacc	2400
gtcctgagtg	tattactttt	caacagatca	aggataattg	cgctaattgag	cttgatctgc	2460
tggcgcagaa	gtattccata	gagcagctga	ccacttactg	gctgcagcca	ggggatgatt	2520
ttgaggaggc	tattagggtta	tatgcaaagg	tggcacttag	gccagattgc	aagtacaaga	2580
tcagcaaact	tgtaaatata	aggaattggt	gctacatttc	tgggaacggg	gccgaggtgg	2640
agatagatac	ggaggatagg	gtggccttta	gatgtagcat	gataaatatg	tggccggggg	2700
tgcttggcat	ggacgggggtg	gttattatga	atgtaagggt	tactggcccc	aatttttagcg	2760
gtacggtttt	cctggccaat	accaacctta	tcctacacgg	tgtaagcttc	tatgggttta	2820
acaatacctg	tgtggaagcc	tggaccgatg	taagggttcg	gggctgtgcc	ttttactgct	2880
gctggaaggg	ggtggtgtgt	cgccccaaaa	gcagggttc	aattaagaaa	tgcctctttg	2940
aaaggtgtac	cttgggtata	ctgtctgagg	gtaactccag	ggtgcgccac	aatgtggcct	3000
ccgactgtgg	ttgcttcatg	ctagtgaaaa	gcgtggctgt	gattaagcat	aacatggtat	3060
gtggcaactg	cgaggacagg	gcctctcaga	tgtgacctg	ctcggacggc	aactgtcacc	3120
tgtgaagac	cattcacgta	gccagccact	ctcgcaaggc	ctggccagtg	tttgagcata	3180
acatactgac	ccgctgttcc	ttgcatttgg	gtaacaggag	gggggtgttc	ctaccttacc	3240
aatgcaatth	gagtcacact	aagatattgc	ttgagcccga	gagcatgtcc	aaggtgaacc	3300
tgaacgggggt	gtttgacatg	accatgaaga	tctggaagggt	gctgaggtac	gatgagaccc	3360
gcaccaggtg	cagaccctgc	gagtgtggcg	gtaaacatat	taggaaccag	cctgtgatgc	3420
tggatgtgac	cgaggagctg	aggcccgatc	acttgggtgct	ggcctgcacc	cgcgctgagt	3480
ttggctctag	cgatgaagat	acagattgag	gtactgaaat	gtgtgggcgt	ggcttaaggg	3540
tgggaaagaa	tatataaggt	gggggtctta	tgtagtthttg	tatctgtthtt	gcagcagccg	3600
ccgccgccat	gagcaccaac	tcgtttgatg	gaagcattgt	gagctcatat	ttgacaacgc	3660
gcatgcccc	atgggcccggg	gtgcgtcaga	atgtgatggg	ctccagcatt	gatggtcgcc	3720
ccgtcctgcc	cgcaaactct	actaccttga	cctacgagac	cgtgtctgga	acgccgttgg	3780
agactgcagc	ctccgccgcc	gcttcagccg	ctgcagccac	cgcccgcggg	attgtgactg	3840
actttgctth	cctgagcccc	cttgcaagca	gtgcagcttc	ccgttcatcc	gcccgcgatg	3900
acaagttgac	ggctctthttg	gcacaattgg	attctthtgac	ccgggaactt	aatgtcgtht	3960
ctcagcagct	gttggatctg	cgccagcagg	thtctgcct	gaaggcttcc	thccctccca	4020

atgcggttta	aaacataaat	aaaaaaccag	actctgtttg	gattttggatc	aagcaagtgt	4080
cttgctgtct	ttatttaggg	gttttgcgcg	cgcggtaggc	ccgggaccag	cggtctcggt	4140
cgttgagggt	cctgtgtatt	ttttccagga	cgtggtaaag	gtgactctgg	atgttcagat	4200
acatgggcat	aagcccgtct	ctgggggtgga	ggtagcacca	ctgcagagct	tcatgctgcg	4260
gggtggtggt	gtagatgatc	cagtcgtagc	aggagcgctg	ggcgtggtgc	ctaaaaatgt	4320
ctttcagtag	caagctgatt	gccaggggca	ggcccttggt	gtaagtgttt	acaaagcggt	4380
taagctggga	tgggtgcata	cgtggggata	tgagatgcat	cttggactgt	attttttaggt	4440
tggctatggt	cccagccata	tccctccggg	gattcatggt	gtgcagaacc	accagcacag	4500
tgtatccggt	gcacttgggga	aatttgtcat	gtagcttaga	aggaaatgcg	tggaagaact	4560
tggagacgcc	cttgtgacct	ccaagatfff	ccatgcattc	gtccataatg	atggcaatgg	4620
gcccacgggc	ggcggcctgg	gcgaagatat	ttctgggatc	actaacgtca	tagttgtggt	4680
ccaggatgag	atcgtcatag	gccatfff	caaagcgcg	gcgagggtg	ccagactgcg	4740
gtataatggt	tccatccggc	ccagggcggt	agttaccctc	acagatttgc	atttcccacg	4800
ctttgagttc	agatgggggg	atcatgtcta	cctgcggggc	gatgaagaaa	acggtttccg	4860
gggtagggga	gatcagctgg	gaagaaagca	ggttcctgag	cagctgcgac	ttaccgcagc	4920
cggtggggcc	gtaaatcaca	cctattaccg	ggtgcaactg	gtagttaaga	gagctgcagc	4980
tgccgtcatc	cctgagcagg	ggggccactt	cgtaagcat	gtccctgact	cgcatgtfff	5040
ccctgaccaa	atccgccaga	aggcgctcgc	cgcccagcga	tagcagttct	tgcaaggaag	5100
caaagttfff	caacggtttg	agaccgtccg	ccgtaggcat	gctfff	gtttgaccaa	5160
gcagttccag	gcggtccac	agctcgggtca	cctgctctac	ggcatctcga	tccagcatat	5220
ctcctcgfff	cgcggttg	ggcggtttc	gctgtacggc	agtagtcggt	gctcgtccag	5280
acgggccagg	gtcatgtctt	tccacggggc	cagggtcctc	gtcagcgtag	tctgggtcac	5340
ggtgaagggg	tgcgctccgg	gctgcgcgct	ggccagggtg	cgcttgaggc	tggtcctgct	5400
ggtgctgaag	cgctgccggt	cttcgccctg	cgcgctcgcc	aggtagcatt	tgaccatggt	5460
gtcatagtcc	agcccctccg	cggcgtggcc	cttggcgcg	agcttgccct	tggaggaggc	5520
gccgcacgag	gggcagtgca	gacttttgag	ggcgtagagc	ttgggcgcga	gaaataccga	5580
ttccggggag	taggcacccg	cgccgcaggc	ccgcagacg	gtctcgcat	ccacgagcca	5640
ggtgagctct	ggccgttcgg	ggtcaaaaac	caggtttccc	ccatgctfff	tgatgcgfff	5700
cttacctctg	gtttccatga	gccgggtgtc	acgctcgggtg	acgaaaaggc	tgtccgtgtc	5760
cccgtataca	gacttgagag	gcctgtcctc	gagcgggtgt	ccgcggtcct	cctcgtatag	5820
aaactcggac	cactctgaga	caaaggctcg	cgtccaggcc	agcacgaagg	aggctaagtg	5880
ggaggggtag	cggtcgttgt	ccactagggg	gtccactcgc	tccagggtgt	gaagacacat	5940

gtcgccctct	tcggcatcaa	ggaaggtgat	tggttttag	gtgtaggcca	cgtgaccggg	6000
tgttcctgaa	ggggggctat	aaaagggggt	gggggcgcgt	tcgtcctcac	tctcttcgcg	6060
atcgctgtct	gcgagggcca	gctggtgggg	tgagtactcc	ctctgaaaag	cgggcatgac	6120
ttctgcgcta	agattgtcag	tttccaaaaa	cgaggaggat	ttgatattca	cctggcccgc	6180
ggtgatgcct	ttgaggggtg	ccgcatccat	ctggtcagaa	aagacaatct	ttttgttgtc	6240
aagcttggtg	gcaaacgacc	cgtagagggc	gttgacagc	aacttggcga	tggagcgag	6300
ggtttggttt	ttgtcgcgat	cggcgcgcgc	cttggccgcg	atgttttagct	gcacgtattc	6360
gcgcgcaacg	caccgccatt	cgggaaagac	ggtggtgcgc	tcgtcgggca	ccaggtgcac	6420
gcgccaaccg	cggttggtgca	gggtgacaag	gtcaacgctg	gtggctacct	ctccgcgtag	6480
gcgctcgttg	gtccagcaga	ggcgcccgcc	cttgcgcgag	cagaatggcg	gtagggggtc	6540
tagctgcgtc	tcgtccgggg	ggtctgcgtc	cacggtaaag	acccggggca	gcaggcgcgc	6600
gtcgaagtag	tctatcttgc	atccttgcaa	gtctagcgcc	tgctgccatg	cgcggggcgc	6660
aagcgcgcgc	tcgtatgggt	tgagtggggg	accccatggc	atgggggtggg	tgagcgcgga	6720
ggcgtagcatg	ccgcaaagt	cgtaaacgta	gaggggctct	ctgagtattc	caagatatgt	6780
agggtagcat	cttccaccgc	ggatgctggc	gcgcacgtaa	tcgtatagtt	cgtgcgaggg	6840
agcgaggagg	tcgggaccga	ggttgctacg	ggcgggctgc	tctgctcgga	agactatctg	6900
cctgaagatg	gcatgtgagt	tggatgatat	ggttgacgc	tggaagacgt	tgaagctggc	6960
gtctgtgaga	cctaccgcgt	cacgcacgaa	ggaggcgtag	gagtcgcgca	gcttggtgac	7020
cagctcggcg	gtgacctgca	cgtctagggc	gcagtagtcc	agggtttcct	tgatgatgtc	7080
atacttatcc	tgtccctttt	ttttccacag	ctcgcggttg	aggacaaact	cttcgcggtc	7140
tttccagtag	tcttggtatg	gaaaccgcgc	ggcctccgaa	cggtaagagc	ctagcatgta	7200
gaactggttg	acggcctggt	aggcgcagca	tcccttttct	acgggtagcg	cgtatgcctg	7260
cgcggccctc	cggagcgagg	tgtgggtgag	cgcaaagggt	tccctgacca	tgactttgag	7320
gtactggtat	ttgaagtcag	tgtcgtcgca	tccgcctgc	tcccagagca	aaaagtccgt	7380
gcgctttttg	gaacgcggat	ttggcagggc	gaaggtgaca	tcgttgaaga	gtatctttcc	7440
cgcgcgaggc	ataaagttgc	gtgtgatgcg	gaagggtccc	ggcacctcgg	aacggttggt	7500
aattacctgg	gcggcgagca	cgatctcgtc	aaagccggtg	atgttggtggc	ccacaatgta	7560
aagttccaag	aagcgcggga	tgcccttgat	ggaaggcaat	tttttaagtt	cctcgtaggt	7620
gagctcttca	ggggagctga	gcccgtgctc	tgaaagggcc	cagtctgcaa	gatgaggggt	7680
ggaagcgacg	aatgagctcc	acaggtcacg	ggccattagc	atttgaggt	ggtcgcgaaa	7740
ggtcctaatac	tggcgacct	tggccatttt	ttctgggggtg	atgcagtaga	aggtaagcgg	7800

gtcttgttcc	cagcgggtccc	atccaaggtt	cgcggttagg	tctcgcgcg	cagtcactag	7860
aggctcatct	ccgccgaact	tcatgaccag	catgaagggc	acgagctgct	tcccaaaggc	7920
ccccatccaa	gtataggtct	ctacatcgta	ggtgacaaag	agacgctcgg	tgcgaggatg	7980
cgagccgatc	gggaagaact	ggatctcccg	ccaccaattg	gaggagtggc	tattgatgtg	8040
gtgaaagtag	aagtccctgc	gacggggccga	acactcgtgc	tggcttttgt	aaaaacgtgc	8100
gcagtactgg	cagcgggtgca	cgggctgtac	atcctgcacg	aggttgacct	gacgaccgcg	8160
cacaaggaag	cagagtggga	atttgagccc	ctcgctggc	gggtttggct	ggtggtcttc	8220
tacttcggct	gcttgtcctt	gaccgtctgg	ctgctcgagg	ggagttagcg	tggatcggac	8280
caccacgccg	cgcgagccca	aagtcagat	gtccgcgcgc	ggcggtcggg	gcttgatgac	8340
aacatcgcg	agatgggagc	tgtccatggt	ctggagctcc	cgcggcgtca	ggtcaggcgg	8400
gagctcctgc	aggtttacct	cgcatagacg	ggtcagggcg	cgggctagat	ccaggtgata	8460
cctaattttcc	aggggctggt	tgggtggcgg	gtcgatggct	tgcaagaggc	cgcatccccg	8520
cggcgcgact	acggtaccgc	gcggcgggcg	gtgggcgcgc	ggggtgtcct	tggatgatgc	8580
atctaaaagc	ggtgacgcgg	gcgagcccc	ggaggtaggg	ggggctccgg	acccgccggg	8640
agagggggca	ggggcacgtc	ggcgccgcgc	gcgggcagga	gctggtgctg	cgcgcgtagg	8700
ttgctggcga	acgcgacgac	gcggcggttg	atctcctgaa	tctgggcct	ctgcgtgaag	8760
acgacggggc	cggtgagctt	gagcctgaaa	gagagttcga	cagaatcaat	ttcgggtgtcg	8820
ttgacggcgg	cctggcgcaa	aatctcctgc	acgtctcctg	agttgtcttg	ataggcgatc	8880
tcggccatga	actgctcgat	ctcttcctcc	tggagatctc	cgcgctccgg	tcgctccacg	8940
gtggcggcga	ggtcgttgga	aatgcggggc	atgagctgcg	agaaggcggt	gaggcctccc	9000
tcgttccaga	cgcggtctga	gaccacgccc	ccttcggcat	cgcgggcgcg	catgaccacc	9060
tgcgcgagat	tgagctccac	gtgccggggc	aagacggcgt	agtttcgcag	gcgctgaaag	9120
aggtagttga	gggtggtggc	ggtgtgttct	gccacgaaga	agtacataac	ccagcgctcg	9180
aacgtggatt	cggtgatata	ccccaaaggc	tcaaggcgct	ccatggcctc	gtagaagtcc	9240
acggcgaagt	tgaaaaactg	ggagttgcgc	gccgacacgg	ttaactcctc	ctccagaaga	9300
cggatgagct	cggcgacagt	gtcgcgcaac	tcgcgctcaa	aggctacagg	ggcctcttct	9360
tcttcttcaa	tctcctcttc	cataaggggc	tcccccttct	cttcttcttg	cggcggtggg	9420
ggagggggga	cacggcgggc	acgacggcgc	accgggaggc	ggtcgacaaa	gcgctcgatc	9480
atctccccgc	ggcgacggcg	catggtctcg	gtgacggcgc	ggccgttctc	gcgggggcgc	9540
agttggaaga	cgccgcccgt	catgtcccgg	ttatgggttg	gcggggggct	gccatgcggc	9600
agggatacgg	cgctaacgat	gcacctcaac	aattgttgtg	taggtactcc	gccgccgagg	9660
gacctgagcg	agtccgcac	gaccggatcg	gaaaacctct	cgagaaaggc	gtctaaccag	9720

tcacagtcgc aaggtaggct gagcacctgt gcgggcggca gcgggcggcg gtcggggttg	9780
tttctggcgg aggtgctgct gatgatgtaa ttaaagtagg cggctctgag acggcggatg	9840
gtcgacagaa gcaccatgtc cttgggtccg gcctgctgaa tgcgcaggcg gtcggccatg	9900
ccccaggctt cgttttgaca tcggcgcagg tctttgtagt agtcttgcat gagcctttct	9960
accggcactt cttcttctcc ttctcttgt cctgcatctc ttgcatctat cgctgcggcg	10020
gcggcggagt ttggccgtag gtggcgccct cttcctccca tgcgtgtgac cccgaagccc	10080
ctcatcggct gaagcagggc taggtcggcg acaacgcgct cggctaatat ggcctgctgc	10140
acctgcgtga gggtagactg gaagtcatcc atgtccacaa agcggtggtg tgcgcccgtg	10200
ttgatggtgt aagtgcagtt ggccataacg gaccagttaa cggctctggtg acccggctgc	10260
gagagctcgg tgtacctgag acgcgagtaa gccctcgagt caaatcgtg gtcgttgcaa	10320
gtccgcacca ggtactggta tcccaccaa aagtgcggcg gcggctggcg gtagaggggc	10380
cagcgtaggg tggccggggc tccggggggcg agatcttcca acataaggcg atgatatccg	10440
tagatgtacc tggacatcca ggtgatgccg gcggcgggtg tggaggcgcg cggaaagtcg	10500
cggacgcggt tccagatgtt gcgcagcggc aaaaagtgt ccatggtcgg gacgctctgg	10560
ccggtcaggc gcgcgcaatc gttgacgctc tagaccgtgc aaaaggagag cctgtaagcg	10620
ggcactcttc cgtggtctgg tggataaatt cgcaagggtg tcatggcgga cgaccggggt	10680
tcgagccccc tatccggccg tccgcctgta tccatgcggt taccgcccgc gtgtcgaacc	10740
cagggtgtgcg acgtcagaca acgggggagt gctccttttg gcttccttcc aggcgcggcg	10800
gctgctgcgc tagctttttt ggccactggc cgcgcgcagc gtaagcgggt aggctggaaa	10860
gcgaaagcat taagtggctc gctccctgta gccggagggt tattttccaa gggttgagtc	10920
gcgggacccc cggttcgagt ctcggaaccg ccggactgcg gcgaacgggg gtttgctcc	10980
ccgtcatgca agaccccgtc tgcaaattcc tccggaaaca gggacgagcc cttttttgc	11040
ttttcccaga tgcacccggt gctgcggcag atgcgcccc ctcctcagca gcggcaagag	11100
caagagcagc ggcagacatg cagggcaccc tcccctcctc ctaccgcgtc aggaggggcg	11160
acatccgcgg ttgacgcggc agcagatggt gattacgaac ccccgcgggc ccgggcccgg	11220
cactacctgg acttgaggga gggcgagggc ctggcgcggc taggagcgcc ctctcctgag	11280
cggtaaccba ggggtgcagct gaagcgtgat acgcgtgagg cgtacgtgcc gcggcagaac	11340
ctgtttcgcg accgcgaggg agaggagccc gaggagatgc gggatcgaaa gttccacgca	11400
gggcgcgagc tgcggcatgg cctgaatcgc gagcggttgc tgcgcgagga ggactttgag	11460
cccgcgcgc gaaccgggat tagtcccgcg cgcgcacacg tggcggccgc cgacctggta	11520
accgcatacg agcagacggt gaaccaggag attaaacttc aaaaagctt taacaaccac	11580

gtgcgtacgc	ttgtggcgcg	cgaggaggtg	gctataggac	tgatgcatct	gtgggacttt	11640
gtaagcgcg	tggagcaaaa	cccaaatagc	aagccgctca	tggcgagct	gttccttata	11700
gtgcagcaca	gcagggacaa	cgaggcattc	agggatgcgc	tgctaaacat	agtagagccc	11760
gagggccgct	ggctgctcga	tttgataaac	atcctgcaga	gcatagtgg	gcaggagcgc	11820
agcttgagcc	tggctgacaa	ggtggccgcc	atcaactatt	ccatgcttag	cctgggcaag	11880
ttttacgccc	gcaagatata	ccatacccct	tacgttccca	tagacaagga	ggtaaagatc	11940
gaggggttct	acatgcgcat	ggcgctgaag	gtgcttacct	tgagcgacga	cctgggcgtt	12000
tatcgcaacg	agcgcatcca	caaggccgtg	agcgtgagcc	ggcggcgcg	gctcagcgac	12060
cgcgagctga	tgcacagcct	gcaaagggcc	ctggctggca	cgggcagcgg	cgatagagag	12120
gccgagtcct	actttgacgc	gggcgctgac	ctgcgctggg	ccccaagccg	acgcgccctg	12180
gaggcagctg	gggccggacc	tgggctggcg	gtggcaccgc	cgcgcgctgg	caacgtcggc	12240
ggcgtggagg	aatatgacga	ggacgatgag	tacgagccag	aggacggcga	gtactaagcg	12300
gtgatgtttc	tgatcagatg	atgcaagacg	caacggaccc	ggcgggtgcg	gcggcgctgc	12360
agagccagcc	gtccggcctt	aactccacgg	acgactggcg	ccaggtcatt	gaccgcatca	12420
tgctcgctgac	tgcgcgcaat	cctgacgcgt	tccggcagca	gccgcaggcc	aaccggctct	12480
ccgcaattct	ggaagcggtg	gtcccggcgc	gcgcaaacc	cacgcacgag	aagggtgctgg	12540
cgatcgtaaa	cgcgctggcc	gaaaacaggg	ccatccggcc	cgacgaggcc	ggcctggtct	12600
acgacgcgct	gcttcagcgc	gtggctcggt	acaacagcgg	caacgtgcag	accaacctgg	12660
accggctggt	gggggatgtg	cgcgaggccg	tggcgagcgc	tgagcgcgcg	cagcagcagg	12720
gcaacctggg	ctccatgggt	gcactaaacg	ccttcctgag	tacacagccc	gccaacgtgc	12780
cgcggggaca	ggaggactac	accaactttg	tgagcgccact	gcggctaatt	gtgactgaga	12840
caccgcaaag	tgaggtgtac	cagtctgggc	cagactattt	tttccagacc	agtagacaag	12900
gcctgcagac	cgtaaacctg	agccaggcct	tcaaaaactt	gcaggggctg	tgggggggtgc	12960
gggctcccac	aggcgaccgc	gcgaccgtgt	ctagcttgct	gacgcccac	tcgcgcctgt	13020
tgctgctgct	aatagcgccc	ttcacggaca	gtggcagcgt	gtcccgggac	acatacctag	13080
gtcacttgct	gacactgtac	cgcgaggcca	taggtcaggc	gcatgtggac	gagcatactt	13140
tccaggagat	tacaagtgtc	agccgcgcgc	tggggcagga	ggacacgggc	agcctggagg	13200
caaccctaaa	ctacctgctg	accaaccggc	ggcagaagat	cccctcgttg	cacagtttaa	13260
acagcgagga	ggagcgcat	ttgcgctacg	tgacgagag	cgtgagcctt	aacctgatgc	13320
gcgacgggg	aacgcccagc	gtggcgctgg	acatgaccgc	gcgcaacatg	gaaccgggca	13380
tgtatgcctc	aaaccggccg	tttatcaacc	gcctaattga	ctacttgcat	cgcgcggccg	13440
ccgtgaaccc	cgagtatttc	accaatgcca	tcttgaaccc	gcactggcta	ccgccccctg	13500

gtttctacac	cggggggattc	gaggtgcccg	agggtaacga	tggattcctc	tgggacgaca	13560
tagacgacag	cgtgttttcc	ccgcaaccgc	agaccctgct	agagttgcaa	cagcgcgagc	13620
aggcagaggc	ggcgcctgca	aaggaaagct	tccgcaggcc	aagcagcttg	tccgatctag	13680
gcgctgcggc	cccgcggtca	gatgctagta	gcccatttcc	aagcttgata	gggtctctta	13740
ccagcactcg	caccacccgc	ccgcgcctgc	tgggcgagga	ggagtaccta	aacaactcgc	13800
tgctgcagcc	gcagcgcgaa	aaaaacctgc	ctccggcatt	tcccaacaac	gggatagaga	13860
gcctagtgga	caagatgagt	agatggaaga	cgtacgcgca	ggagcacagg	gacgtgccag	13920
gcccgcgccc	gcccacccgt	cgtcaaaggc	acgaccgtca	gcgggggtctg	gtgtgggagg	13980
acgatgactc	ggcagacgac	agcagcgtcc	tggatttggg	agggagtggc	aacccgtttg	14040
cgcaccttcg	ccccaggctg	gggagaatgt	tttaaaaaaa	aaaaagcatg	atgcaaaata	14100
aaaaactcac	caaggccatg	gcaccgagcg	ttggttttct	tgtattcccc	ttagtatgcg	14160
gcgcgcggcg	atgtatgagg	aaggctcctcc	tccctcctac	gagagtgtgg	tgagcgcggc	14220
gccagtggcg	gcggcgctgg	gttctccctt	cgatgctccc	ctggacccgc	cgtttgtgcc	14280
tccgcggtac	ctgcggccta	ccgggggggag	aaacagcatc	cgttactctg	agttggcacc	14340
cctattcgac	accacccgtg	tgtacctggt	ggacaacaag	tcaacggatg	tggcatccct	14400
gaactaccag	aacgaccaca	gcaactttct	gaccacggtc	attcaaaaca	atgactacag	14460
cccggggggag	gcaagcacac	agaccatcaa	tcttgacgac	cggtcgcact	ggggcggcga	14520
cctgaaaacc	atcctgcata	ccaacatgcc	aaatgtgaac	gagttcatgt	ttaccaataa	14580
gtttaaggcg	cgggtgatgg	tgtcgcgctt	gcctactaag	gacaatcagg	tggagctgaa	14640
atacgagtgg	gtggagttca	cgctgcccga	gggcaactac	tccgagacca	tgaccataga	14700
ccttatgaac	aacgcgatcg	tggagcacta	cttgaaagtg	ggcagacaga	acggggttct	14760
ggaaagcgac	atcggggtaa	agtttgacac	ccgcaacttc	agactggggg	ttgaccccgt	14820
cactggtctt	gtcatgcctg	gggtatatac	aaacgaagcc	ttccatccag	acatcatttt	14880
gctgccagga	tgcgggggtg	acttcaccca	cagccgcctg	agcaacttgt	tgggcatccg	14940
caagcggcaa	cccttccagg	agggtcttag	gatcacctac	gatgatctgg	agggtggtaa	15000
cattcccgcg	ctgttggatg	tggacgccta	ccaggcgagc	ttgaaagatg	acaccgaaca	15060
gggcgggggt	ggcgcaggcg	gcagcaacag	cagtggcagc	ggcgcggaag	agaactccaa	15120
cgcggcagcc	gcggcaatgc	agccggtgga	ggacatgaac	gatcatgcca	ttcgcggcga	15180
cacctttgcc	acacgggctg	aggagaagcg	cgctgaggcc	gaagcagcgg	ccgaagctgc	15240
cgcccccgct	gcgcaacccg	aggctcgagaa	gcctcagaag	aaaccggtga	tcaaaccctt	15300
gacagaggac	agcaagaaac	gcagttacaa	cctaataagc	aatgacagca	ccttcaccca	15360

gtaccgcagc	tggtaccttg	catacaacta	cggcgaccct	cagaccggaa	tccgctcatg	15420
gaccctgctt	tgcactcctg	acgtaacctg	cggctcggag	caggtctact	ggtcgttgcc	15480
agacatgatg	caagacccccg	tgaccttccg	ctccacgcgc	cagatcagca	actttccggt	15540
ggtgggcgcc	gagctggttg	ccgtgcactc	caagagcttc	tacaacgacc	aggccgtcta	15600
ctcccaactc	atccgccagt	ttacctctct	gaccacagtg	ttcaatcgct	ttcccagaaa	15660
ccagattttg	gcgcgccccg	cagccccccac	catcaccacc	gtcagtgaag	acgttcctgc	15720
tctcacagat	cacgggacgc	taccgctgcg	caacagcatc	ggaggagtcc	agcgagtgc	15780
cattactgac	gccagacgcc	gcacctgccc	ctacgtttac	aaggccctgg	gcatagtctc	15840
gccgcgcgtc	ctatcgagcc	gcactttttg	agcaagcatg	tccatcctta	tatcgcccag	15900
caataacaca	ggctggggcc	tgcgcttccc	aagcaagatg	tttggcgggg	ccaagaagcg	15960
ctccgaccaa	caccacagtgc	gcgtgcgcgg	gcactaccgc	gcgccctggg	gcgcgcacaa	16020
acgcggccgc	actgggcgca	ccaccgtcga	tgacgccatc	gacgcggtgg	tggaggaggc	16080
gcgcaactac	acgcccacgc	cggccaccagt	gtccacagtg	gacgcggcca	ttcagaccgt	16140
ggtgcgcgga	gcccggcgct	atgctaaaat	gaagagacgg	cggaggcgcg	tagcacgtcg	16200
ccaccgccgc	cgacccggca	ctgccgccc	acgcgcggcg	gcggccctgc	ttaaccgcgc	16260
acgtcgcacc	ggccgacggg	cggccatgcg	ggccgctcga	aggctggccg	cgggtattgt	16320
cactgtgccc	cccagggtcca	ggcgacgagc	ggccgcccga	gcagccgagg	ccattagtgc	16380
tatgactcag	ggtcgcaggg	gcaacgtgta	ttgggtgcgc	gactcgggta	gcggcctgcg	16440
cgtgcccgtg	cgcacccgcc	ccccgcgcaa	ctagattgca	agaaaaaact	acttagactc	16500
gtactgttgt	atgtatccag	cggcggcgcc	gcgcaacgaa	gctatgtcca	agcgcaaaat	16560
caaagaagag	atgctccagg	tcacgcgcgc	ggagatctat	ggccccccga	agaaggaaga	16620
gcaggattac	aagccccgaa	agctaaagcg	ggtcaaaaag	aaaaagaaag	atgatgatga	16680
tgaacttgac	gacgaggtgg	aactgctgca	cgctaccgcg	cccaggcgac	gggtacagtg	16740
gaaaggctga	cgcgtaaaac	gtgtttttgc	acccggcacc	accgtagtct	ttacgcccgg	16800
tgagcgctcc	acccgcacct	acaagcgcg	gtatgatgag	gtgtacggcg	acgaggacct	16860
gcttgagcag	gccaacgagc	gcctcgggga	gtttgcctac	ggaaagcggc	ataaggacat	16920
gctggcggtg	ccgctggacg	agggcaaccc	aacacctagc	ctaaagcccg	taacactgca	16980
gcagggtgctg	cccgcgcttg	caccgtccga	agaaaagcgc	ggcctaaagc	gcgagtctgg	17040
tgacttgcca	cccaccgtgc	agctgatgg	acccaagcgc	cagcgactgg	aagatgtctt	17100
ggaaaaaatg	accgtggaac	ctgggctgga	gcccagaggtc	cgcgtcgggc	caatcaagca	17160
ggtggcgccg	ggactgggcg	tgacagccgt	ggacgttcag	ataccacta	ccagtagcac	17220
cagtattgcc	accgccacag	agggcatgga	gacacaaacg	tccccggttg	cctcagcggt	17280

ggcggatgcc	gcggtgcagg	cggtcgctgc	ggccgcgtcc	aagacctcta	cggaggtgca	17340
aacggaccgc	tggatgtttc	gcgtttcagc	ccccggcgcc	ccgcgcggtt	cgaggaagta	17400
cggcgccgcc	agcgcgctac	tgcccgaata	tgcctacat	ccttcattg	cgcctacccc	17460
cggctatcgt	ggctacacct	accgccccag	aagacgagca	actacccgac	gccgaaccac	17520
cactggaacc	cgccgccgcc	gtcgccgtcg	ccagcccgtg	ctggccccga	tttccgtgcg	17580
cagggtggct	cgcgaaaggag	gcaggaccct	ggtgctgcca	acagcgcgct	accaccccag	17640
catcgtttaa	aagccggtct	ttgtggttct	tgcagatatg	gccctcacct	gccgcctccg	17700
tttcccgtg	ccgggattcc	gaggaagaat	gcaccgtagg	aggggcatgg	ccggccacgg	17760
cctgacgggc	ggcatgcgtc	gtgcgcacca	ccggcgggcg	cgcgcgtcgc	accgtcgcat	17820
gcgcggcggt	atcctgcccc	tccttattcc	actgatcgcc	gcggcgattg	gcgccgtgcc	17880
cggaattgca	tccgtggcct	tgaggcgca	gagacactga	ttaaaaacaa	gttgcatgtg	17940
gaaaaatcaa	aataaaaagt	ctggactctc	acgctcgctt	ggtcctgtaa	ctattttgta	18000
gaatggaaga	catcaacttt	gcgtctctgg	ccccgcgaca	cggctcgcg	ccgttcatgg	18060
gaaactggca	agatatcggc	accagcaata	tgagcgggtg	cgccttcagc	tggggctcg	18120
tgtggagcgg	cattaaaaat	ttcggttcca	ccgttaagaa	ctatggcagc	aaggcctgga	18180
acagcagcac	aggccagatg	ctgagggata	agttgaaaga	gcaaaatttc	caacaaaagg	18240
tggtagatgg	cctggcctct	ggcattagcg	gggtggtgga	cctggccaac	caggcagtgc	18300
aaaataagat	taacagtaag	cttgatcccc	gccctcccgt	agaggagcct	ccaccggccg	18360
tggagacagt	gtctccagag	gggcgtggcg	aaaagcgtcc	gcgccccgac	agggaagaaa	18420
ctctggtgac	gcaaatagac	gagcctccct	cgtacgagga	ggcactaaag	caaggcctgc	18480
ccaccacccg	tcccatcgcg	cccatggcta	ccggagtgt	gggccagcac	acaccgtaa	18540
cgctggacct	gcctcccccc	gccgacaccc	agcagaaacc	tgtgctgcca	ggcccgaccg	18600
ccgttggtgt	aaccgcgtct	agccgcgcgt	ccctgcgccg	cgccgccagc	ggtccgcgat	18660
cgttgcggcc	cgtagccagt	ggcaactggc	aaagcacact	gaacagcatc	gtgggtctgg	18720
gggtgcaatc	cctgaagcgc	cgacgatgct	tctgaatagc	taacgtgtcg	tatgtgtgtc	18780
atgtatgcgt	ccatgtcgcc	gccagaggag	ctgctgagcc	gccgcgcgcc	cgctttccaa	18840
gatggctacc	ccttcgatga	tgccgcagtg	gtcttacatg	cacatctcgg	gccaggacgc	18900
ctcggagtac	ctgagccccg	ggctggtgca	gtttgcccgc	gccaccgaga	cgtacttcag	18960
cctgaataac	aagtttagaa	accccacggg	ggcgcctacg	cacgacgtga	ccacagaccg	19020
gtcccagcgt	ttgacgctgc	ggttcacccc	tgtggaccgt	gaggatactg	cgtactcgta	19080
caaggcgcgg	ttcacccctag	ctgtgggtga	taaccgtgtg	ctggacatgg	cttcacgta	19140

ctttgacatc	cgcggcgtgc	tggacagggg	ccctactttt	aagccctact	ctggcactgc	19200
ctacaacgcc	ctggctccca	aggggtgccc	aaatccttgc	gaatgggatg	aagctgctac	19260
tgctcttgaa	ataaacctag	aagaagagga	cgatgacaac	gaagacgaag	tagacgagca	19320
agctgagcag	caaaaaactc	acgtatttgg	gcaggcgcct	tattctggta	taaatattac	19380
aaaggagggt	attcaaatag	gtgtcgaagg	tcaaacacct	aaatatgccg	ataaaacatt	19440
tcaacctgaa	cctcaaatag	gagaatctca	gtggtacgaa	actgaaatta	atcatgcagc	19500
tgggagagtc	cttaaaaaga	ctaccccaat	gaaaccatgt	tacggttcat	atgcaaaacc	19560
cacaaatgaa	aatggagggc	aaggcattct	tgtaaagcaa	caaaatggaa	agctagaaag	19620
tcaagtggaa	atgcaatttt	tctcaactac	tgaggcgacc	gcaggcaatg	gtgataactt	19680
gactcctaaa	gtggtattgt	acagtgaaga	tgtagatata	gaaaccccag	acactcatat	19740
ttcttacatg	cccactatta	aggaaggtaa	ctcacgagaa	ctaattgggc	aacaatctat	19800
gcccacacag	cctaattaca	ttgcttttag	ggacaatttt	attggtctaa	tgtattacaa	19860
cagcacgggt	aatatgggtg	ttctggcggg	ccaagcatcg	cagttgaatg	ctgttgtaga	19920
tttgcaagac	agaaacacag	agctttcata	ccagcttttg	cttgattcca	ttggtgatag	19980
aaccaggtag	ttttctatgt	ggaatcaggc	tgttgacagc	tatgatccag	atgttagaat	20040
tattgaaaat	catggaactg	aagatgaact	tccaaattac	tgctttccac	tgggaggtgt	20100
gattaataca	gagactctta	ccaaggtaaa	acctaaaaca	ggtcaggaaa	atggatggga	20160
aaaagatgct	acagaatttt	cagataaaaa	tgaaataaga	gttggaaata	attttgccat	20220
ggaaatcaat	ctaaatgcca	acctgtggag	aaatttcctg	tactccaaca	tagcgctgta	20280
tttgcccagc	aagctaaagt	acagtccttc	caacgtaaaa	atctctgata	acccaaacac	20340
ctacgactac	atgaacaagc	gagtgggtgg	tcccgggtta	gtggactgct	acattaacct	20400
tggagcacgc	tgggtccctg	actatatgga	caacgtcaac	ccatttaacc	accaccgcaa	20460
tgctggcctg	cgctaccgct	caatgttgct	gggcaatggt	cgctatgtgc	ccttccacat	20520
ccagggtgcct	cagaagttct	ttgccattaa	aaacctcctt	ctcctgccgg	gctcatacac	20580
ctacgagtgg	aacttcagga	aggatgttaa	catggttctg	cagagctccc	taggaaatga	20640
cctaaggggt	gacggagcca	gcattaagtt	tgatagcatt	tgcccttacg	ccaccttctt	20700
ccccatggcc	cacaacaccg	cctccacgct	tgaggccatg	cttagaaacg	acaccaacga	20760
ccagtccttt	aacgactatc	tctccgcgcg	caacatgctc	taccctatac	ccgccaacgc	20820
taccaacgtg	cccatatcca	tcccctcccg	caactgggcg	gctttccgcg	gctgggcctt	20880
cacgcgcctt	aagactaagg	aaaccccatc	actgggctcg	ggctacgacc	cttattacac	20940
ctactctggc	tctataacct	acctagatgg	aaccttttac	ctcaaccaca	cctttaagaa	21000
ggtggccatt	acctttgact	cttctgtcag	ctggcctggc	aatgaccgcc	tgcttacctc	21060

caacgagttt	gaaattaagc	gctcagttga	cggggagggg	tacaacgttg	cccagtgtaa	21120
catgaccaa	gactggttcc	tggtaaaa	gctagcta	tacaacattg	gctaccaggg	21180
cttctatatc	ccagagagct	acaaggaccg	catgtactcc	ttcttttaga	acttccagcc	21240
catgagccgt	caggtggtgg	atgatactaa	atacaaggac	taccaacagg	tgggcatcct	21300
acaccaacac	aacaactctg	gatttggttg	ctaccttgcc	cccaccatgc	gcgaaggaca	21360
ggcctaccct	gctaacttcc	cctatccgct	tataggcaag	accgcagttg	acagcattac	21420
ccagaaaaag	tttctttgcg	atcgcaacct	ttggcgcata	ccatttctca	gtaactttat	21480
gtccatgggc	gcactcacag	acctggggcca	aaaccttctc	tacgccaaact	ccgcccacgc	21540
gctagacatg	acttttgagg	tggatcccat	ggacgagccc	acccttcttt	atgttttggt	21600
tgaagtcttt	gacgtggtcc	gtgtgcaccg	gccgcaccgc	ggcgtcatcg	aaaccgtgta	21660
cctgcgcacg	cccttctcgg	ccggcaacgc	cacaacataa	agaagcaagc	aacatcaaca	21720
acagctgccg	ccatgggctc	cagtgaagcag	gaactgaaag	ccattgtcaa	agatcttggg	21780
tgtggggccat	attttttggg	cacctatgac	aagcgctttc	caggctttgt	ttctccacac	21840
aagctcgcct	gcgccatagt	caatacggcc	ggtcgcgaga	ctggggggcg	acactggatg	21900
gcctttgcct	ggaacccgca	ctcaaaaaca	tgctacctct	ttgagccctt	tggcttttct	21960
gaccagcgac	tcaagcaggt	ttaccagttt	gagtagcag	cactcctgcg	ccgtagcgcc	22020
attgcttctt	cccccgaccg	ctgtataacg	ctggaaaagt	ccacccaaag	cgtacagggg	22080
cccaactcgg	ccgcctgtgg	actattctgc	tgcatgtttc	tccacgcctt	tgccaactgg	22140
ccccaaaactc	ccatggatca	caaccccacc	atgaacctta	ttaccggggg	acccaactcc	22200
atgctcaaca	gtccccaggt	acagcccacc	ctgcgtcgca	accaggaaca	gctctacagc	22260
ttcctggagc	gccactcgcc	ctacttccgc	agccacagtg	cgcagattag	gagcgccact	22320
tctttttgtc	acttgaaaaa	catgtaaaaa	taatgtacta	gagacacttt	caataaaggc	22380
aaatgctttt	atgtgtacac	tctcgggtga	ttatttacct	ccacccttgc	cgtctgcgcc	22440
gtttaaaaat	caaaggggtt	ctgccgcgca	tcgctatgcg	ccactggcag	ggacacgttg	22500
cgatactggg	gtttagtgtc	ccacttaaac	tcaggcacia	ccatccgcgg	cagctcgggtg	22560
aagttttcac	tccacaggtc	gcgcaccatc	accaacgcgt	ttagcaggtc	gggcgccgat	22620
atcttgaagt	cgcagttggg	gcctccgccc	tgccgcgcgc	agttgcgata	cacaggggtg	22680
cagcactgga	acactatcag	cgccgggttg	tgcacgctgg	ccagcacgct	cttgtcggag	22740
atcagatccg	cgtccaggtc	ctccgcgttg	ctcagggcga	acggagtcaa	ctttggtagc	22800
tgccctccca	aaaagggcgc	gtgcccaggc	tttgagttgc	actcgcaccg	tagtggcatc	22860
aaaaggtgac	cgtgcccggt	ctgggcgtta	ggatacagcg	cctgcataaa	agccttgatc	22920

tgcttaaaaag	ccacctgagc	ctttgcgcct	tcagagaaga	acatgccgca	agacttgccg	22980
gaaaactgat	tggccggaca	ggccgcgtcg	tgcacgcagc	accttgcgtc	ggtgttgagg	23040
atctgcacca	catttcggcc	ccaccggttc	ttcacgatct	tggccttgct	agactgctcc	23100
ttcagcgcgc	gctgcccgtt	ttcgctcgtc	acatccattt	caatcacgtg	ctccttattt	23160
atcataatgc	ttccgtgtag	acacttaagc	tcgccttcga	tctcagcgca	gcggtgcagc	23220
cacaacgcgc	agcccgtggg	ctcgatgatc	ttgtagggtca	cctctgcaaa	cgactgcagg	23280
tacgcctgca	ggaatcgccc	catcatcgtc	acaaagggtct	tgttgctggt	gaaggtcagc	23340
tgcaaccgcg	ggtgctcctc	gttcagccag	gtcttgcata	cggccgccag	agcttccact	23400
tggtcaggca	gtagtttgaa	gttcgccttt	agatcgttat	ccacgtggta	cttgtccatc	23460
agcgcgcgcg	cagcctccat	gcccttctcc	cacgcagaca	cgatcggcac	actcagcggg	23520
ttcatcacccg	taatttcact	ttccgcttcg	ctgggctctt	cctcttcctc	ttgcgtccgc	23580
ataccacgcg	ccactgggtc	gtcttcattc	agccgcgcga	ctgtgcgctt	acctcctttg	23640
ccatgcttga	ttagcacccg	tgggttgctg	aaaccaccca	tttgtagcgc	cacatcttct	23700
ctttcttcct	cgctgtccac	gattacctct	ggtgatggcg	ggcgctcggg	cttgggagaa	23760
gggcgcttct	ttttcttctt	gggcgcaatg	gccaaatccg	cgcgcgaggt	cgatggccgc	23820
gggctgggtg	tgcgcggcac	cagcgcgtct	tgtgatgagt	cttcctcgtc	ctcggactcg	23880
atacgccgcc	tcatccgctt	ttttgggggc	gcccggggag	gcggcgggcga	cggggacggg	23940
gacgacacgt	cctccatggt	tgggggacgt	cgcgcgcgac	cgcgtccgcg	ctcggggggtg	24000
gtttcgcgct	gtcctcttct	ccgaactggc	atttccttct	cctataggca	gaaaaagatc	24060
atggagtcag	tcgagaagaa	ggacagccta	accgccccct	ctgagttcgc	caccaccgcc	24120
tccaccgatg	ccgccaacgc	gcctaccacc	ttccccgtcg	aggcaccccc	gcttgaggag	24180
gaggaagtga	ttatcgagca	ggaccaggt	tttghtaagcg	aagacgacga	ggaccgctca	24240
gtaccaacag	aggataaaaa	gcaagaccag	gacaacgcag	aggcaaacga	ggaacaagtc	24300
gggcgggggg	acgaaaggca	tggcgactac	ctagatgtgg	gagacgacgt	gctgttgaag	24360
catctgcagc	gccagtgcgc	cattatctgc	gacgcgttgc	aagagcgagc	cgatgtgccc	24420
ctcgccatag	cggatgtcag	ccttgccctac	gaacgccacc	tattctcacc	gcgcgtaccc	24480
cccaaacgcc	aagaaaacgg	cacatgcgag	cccaaccgcg	gcctcaactt	ctaccccgta	24540
tttgccgtgc	cagaggtgct	tgccacctat	cacatctttt	tccaaaactg	caagataccc	24600
ctatcctgcc	gtgccaacgc	cagccgagcg	gacaagcagc	tggccttgcg	gcagggcgct	24660
gtcatacctg	atatcgccct	gctcaacgaa	gtgccaaaaa	tctttgaggg	tcttgagcgc	24720
gacgagaagc	gcgcggcaaa	cgctctgcaa	caggaaaaca	gcgaaaatga	aagtcactct	24780
ggagtgttgg	tggaaactga	gggtgacaac	gcgcgcctag	ccgtactaaa	acgcagcatc	24840

gaggtcaccc	actttgccta	cccggcactt	aacctacccc	ccaaggtcat	gagcacagtc	24900
atgagtgagc	tgatcgtgcg	ccgtgcgcag	cccctggaga	gggatgcaaa	tttgcaagaa	24960
caaacagagg	agggcctacc	cgcagttggc	gacgagcagc	tagcgcgctg	gcttcaaacg	25020
cgcgagcctg	ccgacttgga	ggagcgcgcg	aaactaatga	tggccgcagt	gctcgttacc	25080
gtggagcttg	agtgcacgca	gcggttcttt	gctgacccgg	agatgcagcg	caagctagag	25140
gaaacattgc	actacacctt	tgcacagggc	tacgtacgcc	aggcctgcaa	gatctccaac	25200
gtggagctct	gcaacctggt	ctcctacctt	ggaattttgc	acgaaaaccg	ccttggggcaa	25260
aacgtgcttc	attccacgct	caagggcgag	gcgcgcgcgc	actacgtccg	cgactgcgtt	25320
tacttatattc	tatgctacac	ctggcagacg	gccatgggcg	tttggcagca	gtgcttggag	25380
gagtgcaccc	tcaaggagct	gcagaaactg	ctaaagcaaa	acttgaagga	cctatggacg	25440
gccttcaacg	agcgctccgt	ggccgcgcac	ctggcggaca	tcattttccc	cgaacgcctg	25500
cttaaaaccc	tgcaacaggg	tctgccagac	ttcaccagtc	aaagcatgtt	gcagaacttt	25560
aggaacttta	tcctagagcg	ctcaggaatc	ttgcccgcca	cctgctgtgc	acttcctagc	25620
gactttgtgc	ccattaagta	ccgcgaatgc	cctccgcgcg	tttggggcca	ctgctacctt	25680
ctgcagctag	ccaactacct	tgcctaccac	tctgacataa	tggaagacgt	gagcgggtgac	25740
ggtctactgg	agtgtcactg	tgcgtgcaac	ctatgcaccc	cgcaccgctc	cctggttttgc	25800
aattcgcagc	tgcttaacga	aagtcaaatt	atcggtagct	ttgagctgca	gggtccctcg	25860
cctgacgaaa	agtccgcggc	tccgggggtt	aaactcactc	cggggctgtg	gacgtcggct	25920
taccttcgca	aatttgtagc	tgaggactac	cacgcccacg	agattaggtt	ctacgaagac	25980
caatcccgcc	cgccaaatgc	ggagcttacc	gcctgcgtca	ttaccagagg	ccacattctt	26040
ggccaattgc	aagccatcaa	caaagccgcg	caagagtttc	tgctacgaaa	gggacggggg	26100
gtttacttgg	acccccagtc	cggcgaggag	ctcaacccaa	tcccccgcc	gccgcagccc	26160
tatcagcagc	agccgcgggc	ccttgcttcc	caggatggca	cccaaaaaga	agctgcagct	26220
gccgccgcca	cccacggacg	aggaggaata	ctgggacagt	caggcagagg	aggttttggg	26280
cgaggaggag	gaggacatga	tggaagactg	ggagagccta	gacgaggaag	cttccgaggt	26340
cgaagagggtg	tcagacgaaa	caccgtcacc	ctcggtcgca	ttccctcgc	cggcgcccca	26400
gaaatcgcca	accggttcca	gcatggctac	aacctccgct	cctcaggcgc	cgccggcact	26460
gcccgttcgc	cgacccaacc	gtagatggga	caccactgga	accagggccg	gtaagtccaa	26520
gcagccgccc	ccgttagccc	aagagcaaca	acagcgccaa	ggctaccgct	catggcgcg	26580
gcacaagaac	gccatagttg	cttgcttgca	agactgtggg	ggcaacatct	ccttcgccc	26640
ccgctttctt	ctctaccatc	acggcggtgg	cttcccccg	aacatcctgc	attactaccg	26700

tcattctctac	agcccatact	gcaccggcgg	cagcggcagc	ggcagcaaca	gcagcggcca	26760
cacagaagca	aaggcgaccg	gatagcaaga	ctctgacaaa	gccaagaaa	tccacagcgg	26820
cggcagcagc	aggaggagga	gcgctgcgtc	tggcgcccaa	cgaaccgcgt	tcgaccgcgc	26880
agcttagaaa	caggatTTTT	cccactctgt	atgctatatt	tcaacagagc	aggggccaag	26940
aacaagagct	gaaaataaaa	aacaggtctc	tgcgatccct	caccgcgagc	tgctgtatc	27000
acaaaagcga	agatcagctt	cggcgcacgc	tggaagacgc	ggaggctctc	ttcagtaa	27060
actgcgcgct	gactcttaag	gactagtttc	gcgccttttc	tcaaatttaa	gcgcgaaaac	27120
tacgtcatct	ccagcggcca	caccgcgcgc	cagcacctgt	cgtcagcgcc	attatgagca	27180
aggaaattcc	cacgccctac	atgtggagtt	accagccaca	aatgggactt	gcggctggag	27240
ctgcccaaga	ctactcaacc	cgaataaact	acatgagcgc	gggacccac	atgatatccc	27300
gggtcaacgg	aatccgcgcc	caccgaaacc	gaattctctt	ggaacaggcg	gctattacca	27360
ccacacctcg	taataacctt	aatccccgta	gttggccgcg	tgccctggtg	taccaggaaa	27420
gtcccgctcc	caccactgtg	gtacttccca	gagacgcccc	ggccgaagtt	cagatgacta	27480
actcaggggc	gcagcttgcg	ggcggttttc	gtcacagggt	gcggtcgccc	gggcagggt	27540
taactcacct	gacaatcaga	gggcgaggta	ttcagctcaa	cgacgagtcg	gtgagctcct	27600
cgtttggtct	ccgtccggac	gggacatttc	agatcggcgg	cgcggccgt	ccttcattca	27660
cgcctcgtca	ggcaatccta	actctgcaga	cctcgtcctc	tgagccgcgc	tctggaggca	27720
ttggaactct	gcaatttatt	gaggagtttg	tgccatcggt	ctactttaac	cccttctcgg	27780
gacctcccg	ccactatccg	gatcaattta	ttcctaactt	tgacgcggta	aaggactcgg	27840
cggacggcta	cgactgaatg	ttaagtggag	aggcagagca	actgcgcctg	aaacacctgg	27900
tccactgtcg	ccgccacaag	tgctttgccc	gcgactccgg	tgagttttgc	tactttgaat	27960
tgcccgagga	tcatatcgag	ggcccggcgc	acggcgtccg	gcttaccgcc	cagggagagc	28020
ttgcccgtag	cctgattcgg	gagtttacc	agcgcgccct	gctagttgag	cgggacaggg	28080
gacctcgtgt	tctcactgtg	atttgcaact	gtcctaacct	tggattacat	caagatcttt	28140
gttgccatct	ctgtgctgag	tataataaat	acagaaatta	aatatactg	gggctcctat	28200
cgccatcctg	taaacgccac	cgtcttcacc	cgcacaagca	aaccaaggcg	aaccttacct	28260
ggtaactttta	acatctctcc	ctctgtgatt	tacaacagtt	tcaaccaga	cggagtgagt	28320
ctacgagaga	acctctccga	gctcagctac	tccatcagaa	aaaacaccac	cctccttacc	28380
tgccgggaac	gtacgagtgc	gtcaccggcc	gctgcaccac	acctaccgcc	tgaccgtaaa	28440
ccagactttt	tccggacaga	cctcaataac	tctgtttacc	agaacaggag	gtgagcttag	28500
aaaaccctta	gggtattagg	ccaaaggcgc	agctactgtg	gggtttatga	acaattcaag	28560
caactctacg	ggctattcta	attcagggtt	ctctagaatc	ggggttgggg	ttattctctg	28620

tcttgtgatt	ctctttattc	ttataactaac	gcttctctgc	ctaaggctcg	ccgcctgctg	28680
tgtgcacatt	tgcatttatt	gtcagctttt	taaacgctgg	ggtcgccacc	caagatgatt	28740
aggtacataa	tcctagggtt	actcaccctt	gcgtcagccc	acggtaccac	ccaaaagggtg	28800
gattttaagg	agccagcctg	taatgttaca	ttcgcagctg	aagctaataa	gtgcaccact	28860
cttataaaat	gcaccacaga	acatgaaaag	ctgcttattc	gccacaaaaa	caaaattggc	28920
aagtatgctg	tttatgctat	ttggcagcca	ggtgacacta	cagagtataa	tgttacagtt	28980
ttccagggtg	aaagtcataa	aacttttatg	tatacttttc	cattttatga	aatgtgcgac	29040
attaccatgt	acatgagcaa	acagtataag	ttgtggcccc	cacaaaattg	tgtggaaaac	29100
actggcactt	tctgctgcac	tgctatgcta	attacagtgc	tcgctttggg	ctgtacccta	29160
ctctatatta	aatacaaaaag	cagacgcagc	tttattgagg	aaaagaaaat	gccttaattt	29220
actaagttac	aaagctaata	tcaccactaa	ctgctttact	cgctgcttgc	aaaacaaatt	29280
caaaaagtta	gcattataat	tagaatagga	tttaaaccct	ccggtcattt	cctgctcaat	29340
accattcccc	tgaacaattg	actctatgtg	ggatatgctc	cagcgctaca	accttgaagt	29400
caggcttcct	ggatgtcagc	atctgacttt	ggccagcacc	tgtcccgcgg	atttgttcca	29460
gtccaactac	agcgacccac	cctaacagag	atgaccaaca	caaccaacgc	ggccgcccgt	29520
accggactta	catctaccac	aaatacaccc	caagtttctg	cctttgtcaa	taactgggat	29580
aacttgggca	tgtggtggtt	ctccatagcg	cttatgtttg	tatgccttat	tattatgtgg	29640
ctcatctgct	gcctaaagcg	caaacgcgcc	cgaccaccca	tctatagtcc	catcattgtg	29700
ctacacccaa	acaatgatgg	aatccataga	ttggacggac	tgaaacacat	gttcttttct	29760
cttacagtat	gattaaatga	gacatgattc	ctcgagtttt	tatattactg	acccttggtg	29820
cgcttttttg	tgcggtgctc	acattggctg	cggtttctca	catcgaagta	gactgcattc	29880
cagccttcac	agtctatttg	ctttacggat	ttgtcaccct	cacgctcatc	tgacgcctca	29940
tcactgtggt	catcgccctt	atccagtgca	ttgactgggt	ctgtgtgcgc	tttgcataac	30000
tcagacacca	tccccagtag	agggacagga	ctatagctga	gcttcttaga	attctttaat	30060
tatgaaattt	actgtgactt	ttctgctgat	tatttgcacc	ctatctgcgt	tttgttcccc	30120
gacctccaag	cctcaaagac	atatatcatg	cagattcact	cgtatatgga	atattccaag	30180
ttgctacaat	gaaaaaagcg	atctttccga	agcctgggta	tatgcaatca	tctctgttat	30240
ggtgttctgc	agtaccatct	tagccctagc	tatatatccc	taccttgaca	ttggctggaa	30300
acgaatagat	gccatgaacc	acccaacttt	ccccgcgccc	gctatgcttc	cactgcaaca	30360
agttgttgcc	ggcggctttg	tcccagccaa	tcagcctcgc	cccacttctc	ccacccccac	30420
tgaaatcagc	tactttaatc	taacaggagg	agatgactga	caccctagat	ctagaaatgg	30480

acggaattat tacagagcag cgctgctag aaagacgcag ggcagcggcc gagcaacagc	30540
gcatgaatca agagctccaa gacatgggta acttgacca gtgcaaaagg ggtatctttt	30600
gtctggtaaa gcaggccaaa gtcacctacg acagtaatac caccggacac cgccttagct	30660
acaagttgcc aaccaagcgt cagaaattgg tggatcatggg gggagaaaag cccattacca	30720
taactcagca ctcggtagaa accgaaggct gcattcactc accttgtcaa ggacctgagg	30780
atctctgcac ccttattaag accctgtgcg gtctcaaaga tcttattccc tttaactaat	30840
aaaaaaaaat aataaagcat cacttactta aaatcagtta gcaaatttct gtccagttta	30900
ttcagcagca cctccttgcc ctctctccag ctctgggtatt gcagcttcct cctggctgca	30960
aactttctcc acaatctaaa tggaatgtca gtttcctcct gttcctgtcc atccgcaccc	31020
actatcttca tgttggttga gatgaagcgc gcaagaccgt ctgaagatac cttcaacccc	31080
gtgtatccat atgacacgga aaccggtcct ccaactgtgc cttttcttac tcctcccttt	31140
gtatccccc atgggtttca agagagtccc cctgggggtac tctctttgcg cctatccgaa	31200
cctctagtta cctccaatgg catgcttgcg ctcaaaatgg gcaacggcct ctctctggac	31260
gaggccggca accttacctc ccaaaatgta accactgtga gccacacctc caaaaaacc	31320
aagtcaaaca taaacctgga aatatctgca cccctcacag ttacctcaga agccctaact	31380
gtggctgccg ccgcacctct aatggctcgc ggcaacacac tcaccatgca atcacaggcc	31440
ccgctaaccg tgcacgactc caaacttagc attgccaccc aaggaccctt cacagtgtca	31500
gaaggaaagc tagccctgca aacatcaggc cccctcacca ccaccgatag cagtaccctt	31560
actatcactg cctcaccccc tctaactact gccactggta gcttgggcat tgacttgaaa	31620
gagcccattt atacacaaaa tggaaaacta ggactaaagt acggggctcc tttgcatgta	31680
acagacgacc taaacacttt gaccgtagca actgggtccag gtgtgactat taataatact	31740
tccttgcaaa ctaaagttac tggagccttg ggttttgatt cacaaggcaa tatgcaactt	31800
aatgtagcag gaggactaag gattgattct caaaacagac gccttatact tgatgttagt	31860
tatccgtttg atgctcaaaa ccaactaaat ctaagactag gacagggccc tctttttata	31920
aactcagccc acaacttgga tattaactac aacaaaggcc tttacttggt tacagcttca	31980
aacaattcca aaaagcttga ggttaaccta agcactgcca aggggttgat gtttgacgct	32040
acagccatag ccattaatgc aggagatggg cttgaatttg gttcacctaa tgcaccaaac	32100
acaaatcccc tcaaaacaaa aattggccat ggcctagaat ttgattcaaa caaggctatg	32160
gttcctaaac taggaactgg ccttagtttt gacagcacag gtgccattac agtaggaaac	32220
aaaaataatg ataagctaac tttgtggacc acaccagctc catctcctaa ctgtagacta	32280
aatgcagaga aagatgctaa actcactttg gtcttaacaa aatgtggcag tcaaatactt	32340
gctacagttt cagttttggc tgttaaaggc agtttggctc caatatctgg aacagttcaa	32400

agtgtcatc	ttattataag	atttgacgaa	aatggagtgc	tactaaacaa	ttccttcctg	32460
gaccagaat	attggaactt	tagaaatgga	gatcttactg	aaggcacagc	ctatacaaac	32520
gctgttggat	ttatgcctaa	cctatcagct	tatccaaaat	ctcacggtaa	aactgccaaa	32580
agtaacattg	tcagtcaagt	ttacttaaac	ggagacaaaa	ctaaacctgt	aacactaacc	32640
attacactaa	acggtacaca	ggaaacagga	gacacaactc	caagtgcata	ctctatgtca	32700
ttttcatggg	actggtctgg	ccacaactac	attaatgaaa	tatttgccac	atcctcttac	32760
actttttcat	acattgccc	agaataaaga	atcgtttgtg	ttatgtttca	acgtgtttat	32820
ttttcaattg	cagaaaat	caagtcat	ttcattcagt	agtatagccc	caccaccaca	32880
tagcttatac	agatcacctg	accttaatca	aactcacaga	accctagtat	tcaacctgcc	32940
acctccctcc	caacacacag	agtacacagt	cctttctccc	cggctggcct	taaaaagcat	33000
catatcatgg	gtaacagaca	tattcttagg	tgttatattc	cacacggttt	cctgtcgagc	33060
caaacgctca	tcagtgat	taataaactc	ccggggcagc	tcacttaagt	tcatgtcgct	33120
gtccagctgc	tgagccacag	gctgctgtcc	aacttgcggt	tgcttaacgg	gcggcgaagg	33180
agaagtccac	gcctacatgg	gggtagagtc	ataatcgtgc	atcaggatag	ggcgggtggtg	33240
ctgcagcagc	gcgcgaataa	actgctgccg	ccgccgctcc	gtcctgcagg	aatacaacat	33300
ggcagtggtc	tcctcagcga	tgattcgac	cgccgcagc	ataaggcgcc	ttgtcctccg	33360
ggcacagcag	cgcaccctga	tctcacttaa	atcagcacag	taactgcagc	acagcaccac	33420
aatattgttc	aaaatcccac	agtgcaaggc	gctgtatcca	aagctcatgg	cggggaccac	33480
agaaccacg	tggccatcat	accacaagcg	caggtagatt	aagtggcgac	ccctcataaa	33540
cacgctggac	ataaacatta	cctcttttgg	catgttgtaa	ttcaccacct	cccgtacca	33600
tataaacctc	tgattaaaca	tggcgccatc	caccaccatc	ctaaaccagc	tggccaaaac	33660
ctgcccgcg	gctatacact	gcaggggaacc	gggactggaa	caatgacagt	ggagagccca	33720
ggactcgtaa	ccatggatca	tcatgctcgt	catgatataca	atgttggcac	aacacaggca	33780
cacgtgcata	cacttcctca	ggattacaag	ctcctcccgc	gttagaacca	tatcccaggg	33840
aacaacccat	tcctgaatca	gcgtaaatcc	cacactgcag	ggaagacctc	gcacgtaact	33900
cacgttgtgc	attgtcaaag	tgttacattc	gggcagcagc	ggatgatcct	ccagtatggt	33960
agcgcgggtt	tctgtctcaa	aaggaggtag	acgatcccta	ctgtacggag	tgcgccgaga	34020
caaccgagat	cgtgttggtc	gtagtgtcat	gccaaatgga	acgccggacg	tagtcatatt	34080
tcctgaagca	aaaccaggtg	cgggcgtgac	aaacagatct	gcgtctccgg	tctcgccgct	34140
tagatcgctc	tgtgtagtag	ttgtagtata	tccactctct	caaagcatcc	aggcgccccc	34200
tggcttcggg	ttctatgtaa	actccttcat	gcgccgctgc	cctgataaca	tccaccaccg	34260

cagaataagc cacacccagc caacctacac attcgttctg cgagtcacac acgggaggag 34320
cggaagagc tggaagaacc atgttttttt ttttattcca aaagattatc caaacctca 34380
aatgaagat ctattaagtg aacgcgctcc cctccggtgg cgtgggtcaa ctctacagcc 34440
aaagaacaga taatggcatt tgtaagatgt tgcacaatgg cttccaaaag gcaaacggcc 34500
ctcacgtcca agtggacgta aaggctaaac ccttcagggg gaatctcctc tataaacatt 34560
ccagcacctt caaccatgcc caaataattc tcatctcgcc accttctcaa tatactctta 34620
agcaaatccc gaatattaag tccggccatt gtaaaaatct gctccagagc gccctccacc 34680
ttcagcctca agcagcgaat catgattgca aaaattcagg ttcctcacag acctgtataa 34740
gattcaaaag cggaacatta acaaaaatac cgcgatcccg taggtccctt cgcagggcca 34800
gctgaacata atcgtgcagg tctgcacgga ccagcgcggc cacttccccg ccaggaacca 34860
tgacaaaaga acccacactg attatgacac gcatactcgg agctatgcta accagcgtag 34920
ccccgatgta agcttggtgc atggggcggc atataaaatg caaggtgctg ctcaaaaaat 34980
caggcaaagc ctgcgcgcaa aaagaaagca catcgtagtc atgctcatgc agataaaggc 35040
aggtaaagtc cggaaccacc acagaaaaag acaccatttt tctctcaaac atgtctgcgg 35100
gtttctgcat aaacacaaaa taaaataaca aaaaaacatt taaacattag aagcctgtct 35160
tacaacagga aaaacaaccc ttataagcat aagacggact acggccatgc cggcgtgacc 35220
gtaaaaaac tggtcaccgt gattaaaaag caccaccgac agctcctcgg tcatgtccgg 35280
agtcataatg taagactcgg taaacacatc aggttgattc acatcgggtca gtgctaaaaa 35340
gcgaccgaaa tagcccgggg gaatacatac ccgcaggcgt agagacaaca ttacagcccc 35400
cataggaggt ataacaaaat taataggaga gaaaaacaca taaacacctg aaaaaccctc 35460
ctgcctaggc aaaatagcac cctcccgtc cagaacaaca tacagcgctt ccacagcggc 35520
agccataaca gtcagcctta ccagtaaaaa agaaaaccta ttaaaaaaac accactcgac 35580
acggcaccag ctcaatcagt cacagtgtaa aaaagggcca agtgcagagc gagtatatat 35640
aggactaaaa aatgacgtaa cggttaaagt ccacaaaaaa caccagaaa accgcacgcg 35700
aacctacgcc cagaaacgaa agccaaaaaa ccacaaactt cctcaaactg tcacttccgt 35760
tttcccacgt tacgtaactt cccattttta gaaaactaca attcccaaca catacaagtt 35820
actccgcctt aaaacctacg tcaccgcgcc cgttcccacg ccccgcgcca cgtcaciaac 35880
tccaccccct cattatcata ttggcttcaa tccaaaataa ggtatattat tgatgatg 35938

<210> 3
<211> 35937
<212> ADN
<213> Adenovirus de tipo 2
<400> 3

catcatcata atatacctta ttttggattg aagccaatat gataatgagg ggggtggagtt	60
tgtgacgtgg cgcggggcgt gggaacgggg cgggtgacgt agtagtgtgg cggaagtgtg	120
atgttgcaag tgtggcggaa cacatgtaag cgccggatgt ggtaaaagtg acgttttttg	180
tgtgcgccgg tgtatacggg aagtgacaat tttcgcgcgg ttttaggcgg atgttgtagt	240
aaatttgggc gtaaccaagt aatgtttggc ctttttcgcg ggaaaactga ataagaggaa	300
gtgaaatctg aataattctg tgttactcat agcgcgtaat atttgtctag ggccgcgggg	360
actttgaccg tttacgtgga gactcgccca ggtgtttttc tcagggtgtt tccgcgttcc	420
gggtcaaagt tggcgtttta ttattatagt cagctgacgc gcagtgtatt tatacccgg	480
gagttcctca agaggccact cttgagtgcc agcgagtaga gttttctcct ccgagccgct	540
ccgacaccgg gactgaaaat gagacatatt atctgccacg gaggtgttat taccgaagaa	600
atggccgcca gtcttttgga ccagctgac gaagaggtag tggctgataa tcttccacct	660
cctagccatt ttgaaccacc tacccttcac gaactgtatg atttagacgt gacggccccc	720
gaagatccca acgaggaggc ggtttcgcag atttttcccg agtctgtaat gttggcgggtg	780
caggaaggga ttgacttatt cacttttcgc ccggcgcccg gttctccgga gccgcctcac	840
ctttcccggc agcccagaca gccggagcag agagccttgg gtccggtttc tatgccaaac	900
cttgtgccgg aggtgatcga tcttacctgc cagcaggctg gctttccacc cagtgcgcac	960
gaggatgaag aggggtgagga gtttgtgtta gattatgtgg agcaccgccg gcacggttgc	1020
aggtcttgtc attatcaccg gaggaatacg ggggaccag atattatgtg ttcgctttgc	1080
tatatgagga cctgtggcat gtttgtctac agtaagtga aattatgggc agtcgggtgat	1140
agagtgggtg gtttgggtgtg gtaatttttt ttttaatttt acagttttgt ggtttaaaga	1200
attttgtatt gtgatttttt aaaaggctct gtgtctgaac ctgagcctga gcccgagcca	1260
gaaccggagc ctgcaagacc taccggcgt cctaaattgg tgcctgctat cctgagacgc	1320
ccgacatcac ctgtgtctag agaatgcaat agtagtacgg atagctgtga ctccggctct	1380
tctaacacac ctctgagat acaccgggtg gtcccgtgt gcccattaa accagttgcc	1440
gtgagagttg gtgggcgtcg ccaggctgtg gaatgtatcg aggacttgct taacgagtct	1500
gggcaacctt tggacttgag ctgtaaacgc ccaggccat aagggtgaaa cctgtgattg	1560
cgtgtgtggt taacgccttt gtttgcgtga tgagttgatg taagtttaat aaagggtgag	1620
ataatgttta acttgcatgg cgtgttaaat ggggcggggc ttaaagggtg tataatgcgc	1680
cgtgggctaa tcttggttac atctgacctc atggaggctt gggagtgttt ggaagatttt	1740
tctgctgtgc gtaacttgct ggaacagagc tctaacagta cctcttggtt ttggagggtt	1800
ctgtggggct cctcccaggc aaagttagtc tgcagaatta aggaggatta caagtgggaa	1860
tttgaagagc ttttgaaatc ctgtggtgag ctgtttgatt ctttgaatct gggtcaccag	1920

gcgccttttcc	aagagaaggt	catcaagact	ttggatTTTT	ccacaccggg	gcgcgctgcg	1980
gctgctgttg	cttttttgag	ttttataaag	gataaatgga	gcgaagaaac	ccatctgagc	2040
ggggggtacc	tgctggattt	tctggccatg	catctgtgga	gagcgggtgg	gagacacaag	2100
aatcgctgc	tactgttgtc	ttccgtccgc	ccggcaataa	taccgacgga	ggagcaacag	2160
caggaggaag	ccaggcggcg	gcggcggcag	gagcagagcc	catggaaccc	gagagccggc	2220
ctggaccctc	gggaatgaat	gttgtagcag	tggtgaact	gtttccagaa	ctgagacgca	2280
ttttaaccat	taacgaggat	gggcaggggc	taaaggggt	aaagaggag	cggggggctt	2340
ctgaggctac	agaggaggct	aggaatctaa	cttttagctt	aatgaccaga	caccgtcctg	2400
agtgtgttac	ttttcagcag	attaaggata	attgcgctaa	tgagcttgat	ctgctggcgc	2460
agaagtattc	catagagcag	ctgaccactt	actggctgca	gccaggggat	gattttgagg	2520
aggctattag	ggtatatgca	aaggtggcac	ttaggccaga	ttgcaagtac	aagattagca	2580
aacttgtaaa	tatcaggaat	tgttgctaca	tttctgggaa	cggggcccag	gtggagatag	2640
atacgagga	taggggtggc	tttagatgta	gcatgataaa	tatgtggccg	ggggtgcttg	2700
gcatggacgg	ggtggttatt	atgaatgtga	ggtttactgg	tcccaatttt	agcggtagcg	2760
ttttcctggc	caataccaat	cttatcctac	acgggtgaag	cttctatggg	tttaacaata	2820
cctgtgtgga	agcctggacc	gatgtaaggg	ttcggggctg	tgccctttac	tgctgctgga	2880
aggggggtgg	gtgtgcggcc	aaaagcaggg	cttcaattaa	gaaatgcctg	tttgaaaggt	2940
gtaccttggg	tatcctgtct	gagggttaact	ccagggtgcg	ccacaatgtg	gcctccgact	3000
gtggttgctt	catgctagtg	aaaagcgtgg	ctgtgattaa	gcataacatg	gtgtgtggca	3060
actgcgagga	cagggcctct	cagatgctga	cctgctcgga	cggcaactgt	cacttgctga	3120
agaccattca	cgtagccagc	cactctcgca	aggcctggcc	agtgtttgag	cacaacatac	3180
tgaccgcgtg	ttccttgcat	ttgggtaaca	ggaggggggt	gttcctacct	taccaatgca	3240
atgtgagtca	cactaagata	ttgcttgagc	ccgagagcat	gtccaagggt	aacctgaacg	3300
gggtgtttga	catgaccatg	aagatctgga	aggtgctgag	gtacgatgag	acccgcacca	3360
ggtgcagacc	ctgcgagtgt	ggcggtaaac	atattaggaa	ccagcctgtg	atgctggatg	3420
tgaccgagga	gctgaggccc	gatcacttgg	tgctggcctg	caccgcgct	gagtttggct	3480
ctagcgatga	agatacagat	tgaggtagct	aaatgtgtgg	gcgtggctta	aggggtgggaa	3540
agaatatata	aggtgggggt	ctcatgtagt	tttgtatctg	ttttgcagca	gccgccgcca	3600
tgagcgccaa	ctcgtttgat	ggaagcattg	tgagctcata	tttgacaacg	cgcatgcccc	3660
catgggcccg	ggtgcgtcag	aatgtgatgg	gctccagcat	tgatggtcgc	cccgtcctgc	3720
ccgcaaactc	tactaccttg	acctacgaga	ccgtgtctgg	aacgccgttg	gagactgcag	3780

cctccgccgc	cgcttcagcc	gctgcagcca	ccgccgcggg	gattgtgact	gactttgctt	3840
tcctgagccc	gcttgcaagc	agtgcagctt	cccgttcac	cgcccgcat	gacaagttga	3900
cggctctttt	ggcacaattg	gattctttga	ccggggaact	taatgtcgtt	tctcagcagc	3960
tggttgatct	gcgccagcag	gtttctgccc	tgaaggcttc	ctcccctccc	aatgcggttt	4020
aaaacataaa	taaaaaccag	actctgtttg	gattttgatc	aagcaagtgt	cttgctgtct	4080
ttatttaggg	gttttgccgc	cgcggtaggc	ccgggaccag	cggtctcggt	cgttgagggg	4140
cctgtgtatt	ttttccagga	cgtggtaaag	gtgactctgg	atgttcagat	acatgggcat	4200
aagcccgtct	ctgggggtga	ggtagcacca	ctgcagagct	tcatgctgcg	gggtgggtgt	4260
gtagatgatc	cagtcgtagc	aggagcgtg	ggcgtgggtg	ctaaaaatgt	ctttcagtag	4320
caagctgatt	gccaggggca	ggcccttggt	gtaagtgttt	acaaagcggg	taagctggga	4380
tgggtgcata	cgtggggata	tgagatgcat	cttggactgt	attttttaggt	tggctatgtt	4440
cccagccata	tccctccggg	gattcatgtt	gtgcagaacc	accagcacag	tgtatccggg	4500
gcacttgga	aatttgctcat	gtagcttaga	aggaaatgcg	tggaagaact	tggagacgcc	4560
cttgtgacct	ccgagatttt	ccatgcattc	gtccataatg	atggcaatgg	gccacgggc	4620
ggcggcctgg	gcgaagatat	ttctgggatc	actaacgtca	tagttgtgtt	ccaggatgag	4680
atcgtcatag	gccattttta	caaagcgcgg	gcggagggtg	ccagactgcg	gtataatggt	4740
tccatccggc	ccaggggcgt	agttaccctc	acagattttg	atttccacg	ctttgagttc	4800
agatgggggg	atcatgtcta	cctgcggggc	gatgaagaaa	accgtttccg	gggtagggga	4860
gatcagctgg	gaagaaagca	ggttcctgag	cagctgcgac	ttaccgcagc	cggtgggccc	4920
gtaaatcaca	cctattaccg	gctgcaactg	gtagttaaga	gagctgcagc	tgccgtcatc	4980
cctgagcagg	ggggccactt	cgttaagcat	gtccctgact	tgcatgtttt	ccctgaccaa	5040
atgcgccaga	aggcgtctgc	cgccagcga	tagcagttct	tgcaaggaag	caaagttttt	5100
caacggtttg	aggccgtccg	ccgtaggcat	gcttttgagc	gtttgaccaa	gcagttccag	5160
gcggtcccac	agctcgggtca	cgtgctctac	ggcatctcga	tccagcatat	ctcctcgttt	5220
cgcgggttgg	ggcggctttc	gctgtacggc	agtagtcggg	gctcgtccag	acgggcccag	5280
gtcatgtctt	tccacggggc	cagggtcctc	gtcagcgtag	tctgggtcac	ggtgaagggg	5340
tgcgctccgg	gctgcgcgct	ggccagggtg	cgcttgaggc	tggctctgct	ggtgctgaag	5400
cgtgcgccgt	cttcgccctg	cgcgtcggcc	aggtagcatt	tgaccatggt	gtcatagtcc	5460
agcccctccg	cggcgtggcc	cttggcgcgc	agcttgccct	tggaggaggc	gccgcacgag	5520
gggcagtgca	gacttttaag	ggcgtagagc	ttgggcgcga	gaaataccga	ttccggggag	5580
taggcatccg	cgccgcaggc	cccgcagacg	gtctcgcat	ccacgagcca	ggtgagctct	5640
ggccgttcgg	ggtcaaaaac	caggtttccc	ccatgctttt	tgatgcgttt	cttacctctg	5700

gtttccatga gccggtgtcc acgctcgggtg acgaaaaggc tgtccgtgtc cccgtataca	5760
gacttgagag gcctgtcctc gagcgggtgtt ccgcgggtcct cctcgtatag aaactcggac	5820
cactctgaga cgaaggctcg cgtccaggcc agcacgaagg aggctaagtg ggaggggtag	5880
cggtcgttgt ccactagggg gtccactcgc tccagggtgt gaagacacat gtcgccctct	5940
tcggcatcaa ggaaggatgat tggtttatag gtgtaggcca cgtgaccggg tgttcctgaa	6000
ggggggctat aaaagggggg gggggcgcggt tcgtcctcac tctcttcgcg atcgtgtct	6060
gcgagggcca gctgttgggg tgagtactcc ctctcaaaag cgggcatgac ttctgcgcta	6120
agattgtcag tttccaaaaa cgaggaggat ttgatattca cctggcccgc ggtgatgcct	6180
ttgaggggtg ccgcgtccat ctggtcagaa aagacaatct ttttgttgtc aagcttgggtg	6240
gcaaacgacc cgtagagggc gttggacagc aacttggcga tggagcgcag ggtttggttt	6300
ttgtcgcgat cggcgcgctc cttggccgcg atgtttagct gcacgtattc gcgcgcaacg	6360
caccgccatt cgggaaagac ggtggtgcgc tcgtcgggca ccagggtcac gcgccaaccg	6420
cggttgtgca gggtgacaag gtcaacgctg gtggctacct ctccgcgtag gcgctcgttg	6480
gtccagcaga ggcggccgcc cttgcgcgaa cagaatggcg gtagtgggtc tagctgcgtc	6540
tcgtccgggg ggtctgcgtc cacggtaaag accccgggca gcaggcgcgc gtcgaagtag	6600
tctatcttgc atccttgcaa gtctagcgcc tgctgccatg cgcgggcggc aagcgcgcgc	6660
tcgtatgggt tgagtggggg accccatggc atgggggtggg tgagcgcgga ggcgtacatg	6720
ccgcaaagtgt cgtaaacgta gaggggctct ctgagtattc caagatatgt agggtagcat	6780
cttccaccgc ggatgctggc gcgcacgtaa tcgtatagtt cgtgcgaggg agcgaggagg	6840
tcgggaccga ggttgctacg ggcgggctgc tctgctcgga agactatctg cctgaagatg	6900
gcatgtgagt tggatgatat ggttggacgc tggaagacgt tgaagctggc gtctgtgaga	6960
cctaccgcgt cacgcacgaa ggaggcgtag gagtgcgcga gcttgttgac cagctcggcg	7020
gtgacctgca cgtctagggc gcagtagtcc agggtttcct tgatgatgtc atacttatcc	7080
tgtccctttt ttttccacag ctgcgggttg aggacaaact cttcgcggtc tttccagtac	7140
tcttgatcg gaaaccgcgc ggcctccgaa cggtaagagc ctagcatgta gaactggttg	7200
acggcctggg aggcgcagca tcccttttct acgggtagcg cgtatgcctg cgcggccttc	7260
cggagcaggg tgtgggtgag cgcaaagggtg tccctaacca tgactttgag gtactggtat	7320
ttgaagtcag tgtcgtcgca tccgccctgc tcccagagca aaaagtccgt gcgctttttg	7380
gaacgcgggt ttggcagggc gaagggtgaca tcgttgaaaa gtatctttcc cgcgcgaggc	7440
ataaagttgc gtgtgatgcg gaagggtccc ggcacctcgg aacggttggt aattacctgg	7500
gcggcgagca cgatctcgtc gaagccgttg atgttgtggc ccacgatgta aagttccaag	7560

aagcgcgggg	tgcccttgat	ggagggcaat	tttttaagtt	cctcgtaggt	gagctcctca	7620
ggggagctga	gcccggttc	tgacagggcc	cagtctgcaa	gatgaggggt	ggaagcgacg	7680
aatgagctcc	acaggtcacg	ggccattagc	atttgcaggt	ggtcgcgaaa	ggtcctaaac	7740
tggcgacct	tggccatttt	ttctgggggtg	atgcagtaga	aggtaagcgg	gtcttggtcc	7800
cagcgggtccc	atccaagggtc	cacggctagg	tctcgcgcgg	cggtcaccag	aggctcatct	7860
ccgccgaact	tcataaccag	catgaagggc	acgagctgct	tcccaaaggc	ccccatccaa	7920
gtataggtct	ctacatcgta	ggtgacaaaag	agacgctcgg	tgcgaggatg	cgagccgatc	7980
gggaagaact	ggatctcccg	ccaccagttg	gaggagtggc	tgttgatgtg	gtgaaagtag	8040
aagtccctgc	gacggggccga	acactcgtgc	tggcttttgt	aaaaacgtgc	gcagtactgg	8100
cagcgggtgca	cgggctgtac	atcctgcacg	aggttgacct	gacgaccgcg	cacaaggaag	8160
cagagtggga	atttgagccc	ctcgcctggc	gggtttggct	ggtggtcttc	tacttcggct	8220
gcttgctcctt	gaccgtctgg	ctgctcgagg	ggagttatgg	tggatcggac	caccacgccg	8280
cgcgagccca	aagtccagat	gtccgcgcgc	ggcggtcgga	gcttgatgac	aacatcgcgc	8340
agatgggagc	tgtccatggt	ctggagctcc	cgcggcgaca	ggtcaggcgg	gagctcctgc	8400
aggtttacct	cgcatagccg	ggtcagggcg	cgggctaggt	ccaggtgata	cctgatttcc	8460
aggggctggg	tgggtggcggc	gtcgatgact	tgcaagaggc	cgcacccccg	cggcgcgact	8520
acggtaccgc	gcggcgggcg	gtgggcccgcg	gggggtgtcct	tggatgatgc	atctaaaagc	8580
ggtgacgcgg	gcggggcccc	ggaggtaggg	ggggctcggg	acccgcgggg	agagggggca	8640
ggggcacgtc	ggcgccgcgc	gcgggcagga	gctggtgctg	cgcgcggagg	ttgctggcga	8700
acgcgacgac	gcggcggttg	atctcctgaa	tctggcgcct	ctgcgtgaag	acgacgggcc	8760
cgggtgagctt	gaacctgaaa	gagagttcga	cagaatcaat	ttcgggtgtcg	ttgacggcgg	8820
cctggcgcaa	aatctcctgc	acgtctcctg	agttgtcttg	ataggcgatt	tcggccatga	8880
actgctcgat	ctcttcctcc	tggagatctc	cgcgtccggc	tcgctccacg	gtggcggcga	8940
ggtcgttgga	gatgcgggcc	atgagctgcg	agaaggcggt	gaggcctccc	tcgttccaga	9000
cgcggctgta	gaccacgccc	ccttcggcat	cgcgggcgcg	catgaccacc	tgcgcgagat	9060
tgagctccac	gtgccggggcg	aagacggcgt	agtttcgcag	gcgctgaaag	aggtagttga	9120
gggtggtggc	ggtgtgttct	gccacgaaga	agtacataac	ccagcgtcgc	aacgtggatt	9180
cgttgatatc	ccccaaggcc	tcaaggcgct	ccatggcctc	gtagaagtcc	acggcgaagt	9240
tgaaaaactg	ggagttgcgc	gccgacacgg	ttaactcctc	ctccagaaga	cggatgagct	9300
cggcgacagt	gtcgcgcacc	tgcgcgtcaa	aggctacagg	ggcctcttct	tcttcaatct	9360
cctcttccat	aagggcctcc	ccttcttctt	cttcttctgg	cggcgggtggg	ggagggggga	9420
cacggcgggcg	acgacggcgc	accgggaggg	ggtcgacaaa	gcgctcgatc	atctccccgc	9480

ggcgacggcg	catggtctcg	gtgacggcgc	ggccgttctc	gcggggggcg	agttggaaga	9540
cgccgcccgt	catgtcccgg	ttatgggttg	gcggggggct	gccgtgcggc	agggatacgg	9600
cgctaacgat	gcatctcaac	aattgttgtg	taggtactcc	gccaccgagg	gacctgagcg	9660
agtccgcata	gaccggatcg	gaaaacctct	cgagaaaggc	gtctaaccag	tcacagtcgc	9720
aaggtaggct	gagcaccgtg	gcggggcggc	gcgggtggcg	gtcgggggtg	tttctggcgg	9780
aggtgctgct	gatgatgtaa	ttaaagtagg	cggtcttgag	acggcggatg	gtcgacagaa	9840
gcaccatgtc	cttgggtccg	gcctgctgaa	tgcgcaggcg	gtcggccatg	ccccaggctt	9900
cgttttgaca	tcggcgcagg	tctttgtagt	agtcttgcat	gagcctttct	accggcactt	9960
cttcttctcc	ttcctcttgt	cctgcatctc	ttgcatctat	cgctacggcg	gcggcggagt	10020
ttggccgtag	gtggcgccct	cttcctccca	tgcgtgtgac	cccgaagccc	ctcatcggtc	10080
gaagcagggc	caggtcggcg	acaacgcgct	cggctaatat	ggcctgctgc	acctgcgtga	10140
gggtagactg	gaagtcatac	atgtccacaa	agcggtggtg	tgcgcccgtg	ttgatgggtg	10200
aagtgcagtt	ggccataacg	gaccagttaa	cggtctgggt	acccggctgc	gagagctcgg	10260
tgtacctgag	acgcgagtaa	gcccttgagt	caaagacgta	gtcgttgcaa	gtccgcacca	10320
ggtactgata	tcccaccaaa	aagtgcggcg	gcggctggcg	gtagaggggc	cagcgtaggg	10380
tggccggggc	tccggggggc	aggtcttcca	acataaggcg	atgatatccg	tagatgtacc	10440
tggacatcca	ggtgatgccg	gcggcggttg	tggaggcgcg	cggaaagtcg	cggacgcggt	10500
tccagatggt	gcgcagcggc	aaaaagtgct	ccatggctcg	gacgctctgg	ccggtgaggg	10560
gtgcgcagtc	gttgacgctc	tagaccgtgc	aaaaggagag	cctgtaagcg	ggcactcttc	10620
cgtggtctgg	tggataaatt	cgcaagggtg	tcatggcgga	cgaccggggg	tcgaaccccg	10680
gatccggccg	tccgccgtga	tccatgcggt	taccgcccgc	gtgtcgaacc	cagggtgtgcg	10740
acgtcagaca	acggggggagc	gctccttttg	gcttccttcc	aggcgcgggc	gctgctgcgc	10800
tagctttttt	ggccactggc	cgcgcgcggc	gtaagcggtt	aggctggaaa	gcgaaagcat	10860
taagtggctc	gctccctgta	gccggagggt	tattttccaa	gggttgagtc	gcaggacccc	10920
cggttcgagt	ctcggggccg	ccggactgcg	gcgaacgggg	gtttgcctcc	ccgtcatgca	10980
agaccccgct	tgcaaattcc	tccggaaaca	gggacgagcc	ccttttttgc	ttttcccaga	11040
tgcatccggt	gctgcggcag	atgcgcccc	ctcctcagca	gcggcaagag	caagagcagc	11100
ggcagacatg	cagggcaccc	tccccttctc	ctaccgcgtc	aggaggggca	acatccgcgg	11160
ctgacgcggc	ggcagatggt	gattacgaac	ccccgcggcg	ccgggcccgg	cactacctgg	11220
acttgaggga	gggcgagggc	ctggcgcggc	taggagcgcc	ctctcctgag	cgacacccaa	11280
gggtgcagct	gaagcgtgac	acgcgcgagg	cgtacgtgcc	gcggcagaac	ctgtttcgcg	11340

accgcgaggg	agaggagccc	gaggagatgc	gggatcgaaa	gttccacgca	gggcgcgagt	11400
tgcggcatgg	cctgaaccgc	gagcggttgc	tgcgcgagga	ggactttgag	cccgcgcgc	11460
ggaccgggat	tagtcccgcg	cgcgcacacg	tggcgggcgc	cgacctggtg	accgcgtacg	11520
agcagacggt	gaaccaggag	attaactttc	aaaaaagctt	taacaaccac	gtgcgcacgc	11580
ttgtggcgcg	cgaggagggtg	gctataggac	tgatgcatct	gtgggacttt	gtaagcgcg	11640
tggagcaaaa	cccaaatagc	aagccgctca	tggcgcgagct	gttccttata	gtgcagcaca	11700
gcagggacaa	cgaggcattc	agggatgcgc	tgctaaacat	agtagagccc	gagggccgct	11760
ggctgctcga	tttgataaac	attctgcaga	gcatagtggg	gcaggagcgc	agcttgagcc	11820
tggctgacaa	ggtggccgcc	attaactatt	ccatgctcag	tctgggcaag	ttttacgccc	11880
gcaagatata	ccatacccct	tacgttccca	tagacaagga	ggtaaagatc	gaggggttct	11940
acatgcgcat	ggcgttgaag	gtgcttacct	tgagcgacga	cctgggcggt	tatcgcaacg	12000
agcgcatcca	caaggccgtg	agcgtgagcc	ggcgggcgga	gctcagcgac	cgcgagctga	12060
tgcacagcct	gcaaagggcc	ctggctggca	cgggcagcgg	cgatagagag	gccgagtcct	12120
actttgacgc	gggcgctgac	ctgcgctggg	ccccaaagccg	acgcgccctg	gaggcagctg	12180
gggccggacc	tgggctggcg	gtggcaccgc	cgcgcgctgg	caacgtcggc	ggcgtggagg	12240
aatatgacga	ggacgatgag	tacgagccag	aggacggcga	gtactaagcg	gtgatgtttc	12300
tgatcagatg	atgcaagacg	caacggaccc	ggcgggtgcgg	gcgggcgctgc	agagccagcc	12360
gtccggcctt	aactccacgg	acgactggcg	ccagggtcatg	gaccgcatca	tgtcgctgac	12420
tgcgcgtaac	cctgacgcgt	tccggcagca	gccgcaggcc	aaccggctct	ccgcaattct	12480
ggaagcggtg	gtcccggcgc	gcgcaaacc	cacgcacgag	aaggtgctgg	cgatcgtaaa	12540
cgcgctggcc	gaaaacaggg	ccatccggcc	cgatgaggcc	ggcctggtct	acgacgcgct	12600
gcttcagcgc	gtggctcggt	acaacagcgg	caacgtgcag	accaacctgg	accggctggt	12660
gggggatgtg	cgcgaggccg	tggcgcgagc	tgagcgcgcg	cagcagcagg	gcaacctggg	12720
ctccatggtt	gcactaaacg	ccttcctgag	tacacagccc	gccaacgtgc	cgcggggaca	12780
ggaggactac	accaactttg	tgagcgcact	gcggctaata	gtgactgaga	caccgcaaag	12840
tgagggtgtac	cagtccgggg	cagactatct	tttccagacc	agtagacaag	gcctgcagac	12900
cgtaaacctg	agccaggctt	tcaagaactt	gcaggggctg	tgggggggtgc	gggctcccac	12960
aggcgaccgc	gcgaccgtgt	ctagcttgct	gacgcccac	tcgcgcctgt	tgctgctgct	13020
aatagcggcc	ttcacggaca	gtggcagcgt	gtcccgggac	acatacctag	gtcacttgct	13080
gacactgtac	cgcgaggcca	taggtcaggc	gcatgtggac	gagcatactt	tccaggagat	13140
tacaagtgtc	agccgcgcgc	tggggcagga	ggacacgggc	agcctggagg	caacctgaa	13200
ctacctgctg	accaaccggc	ggcagaagat	cccctcgttg	cacagtttaa	acagcgagga	13260

ggagcgcac	ttgcgctatg	tgcagcagag	cgtgagcctt	aacctgatgc	gcgacggggt	13320
aacgcccagc	gtggcgctgg	acatgaccgc	gcgcaacatg	gaaccgggca	tgtatgcctc	13380
aaaccggccg	tttatcaatc	gcctaattga	ctacttgcat	cgcgcggccg	ccgtgaaccc	13440
cgagtatttc	accaatgcca	tcttgaaccc	gcactggcta	ccgccccctg	gtttctacac	13500
cgggggattt	gaggtgcccg	agggtaacga	tggattcctc	tgggacgaca	tagacgacag	13560
cgtgttttcc	ccgcaaccgc	agaccctgct	agagttgcaa	cagcgcgagc	aggcagaggc	13620
ggcgctgcga	aaggaaagct	tccgcaggcc	aagcagcttg	tccgatctag	gcgctgcggc	13680
cccgcggtca	gatgcgagta	gcccatttcc	aagcttgata	gggtctttta	ccagcactcg	13740
caccacccgc	ccgcgcctgc	tgggcgagga	ggagtaccta	aacaactcgc	tgctgcagcc	13800
gcagcgcgaa	aagaacctgc	ctccggcatt	tcccaacaac	gggatagaga	gcctagtgga	13860
caagatgagt	agatggaaga	cgtatgcgca	ggagcacagg	gatgtgcccg	gcccgcgccc	13920
gcccacccgt	cgtcaaaggc	acgaccgtca	gcgggggtctg	gtgtgggagg	acgatgactc	13980
ggcagacgac	agcagcgtcc	tggatttggg	agggagtggc	aaccgtttg	cgcaccttcg	14040
ccccaggctg	gggagaatgt	tttaaaaaaa	aaaaaaaaaa	gcatgatgca	aaataaaaaa	14100
ctcaccaagg	ccatggcacc	gagcgttggg	tttcttgtat	tccccttagt	atgcagcgcg	14160
cggcgatgta	tgaggaagg	cctcctccct	cctacgagag	cgtggtgagc	gcggcgccag	14220
tggcggcggc	gctgggttcc	cccttcgatg	ctcccctgga	cccgcggtt	gtgcctccgc	14280
ggtacctgcg	gcctaccggg	gggagaaaca	gcatccgtta	ctctgagttg	gcacccctat	14340
tcgacaccac	ccgtgtgtac	cttgtggaca	acaagtcaac	ggatgtggca	tcctgaact	14400
accagaacga	ccacagcaac	tttctaacca	cggtcattca	aaacaatgac	tacagcccgg	14460
gggaggcaag	cacacagacc	atcaatcttg	acgaccgttc	gcactggggc	ggcgacctga	14520
aaaccatcct	gcataccaac	atgccaaatg	tgaacgagtt	catgtttacc	aataagttta	14580
aggcgcgggt	gatggtgtcg	cgtctgctta	ctaaggacaa	acaggtggag	ctgaaatatg	14640
agtgggtgga	gttcacgctg	cccaggggca	actactccga	gaccatgacc	atagacctta	14700
tgaacaacgc	gatcgtggag	cactacttga	aagtgggcag	gcagaacggg	gttctggaaa	14760
gcgacatcgg	ggtaaagttt	gacacccgca	acttcagact	ggggtttgac	ccagtacttg	14820
gtcttgtcat	gcctggggta	tatacaaacg	aagccttcca	tccagacatc	atthtctgctg	14880
caggatgcgg	ggtggacttc	acccacagcc	gcctgagcaa	cttggtgggc	atccgcaagc	14940
ggcaaccctt	ccaggagggc	tttaggatca	cctacgatga	cctggagggt	ggtaacattc	15000
ccgcactggt	ggatgtggac	gcctaccagg	caagcttaaa	agatgacacc	gaacagggag	15060
gggatggcgc	aggcggcggc	aacaacagtg	gcagcggcgc	ggaagagaac	tccaacgcgg	15120

cagccgcggc	aatgcagccg	gtggaggaca	tgaacgatca	tgccattcgc	ggcgacacct	15180
ttgccacacg	ggcggaggag	aagcgcgctg	aggccgaggc	agcggcagaa	gctgccgccc	15240
ccgctgcgca	acccgaggtc	gagaagcctc	agaagaaacc	ggtgatcaaa	cccctgacag	15300
aggacagcaa	gaaacgcagt	tacaacctaa	taagcaatga	cagcaccttc	accagtacc	15360
gcagctggta	ccttgcatat	aactacggcg	accctcagac	cgggatccgc	tcatggaccc	15420
tcctttgcac	tcctgacgta	acctgcggct	cggagcaggt	ctactggtcg	ttgccagaca	15480
tgatgcaaga	ccccgtgacc	ttccgctcca	cgagccagat	cagcaacttt	ccggtggtgg	15540
gcgccgagct	gttgcgcgtg	cactccaaga	gcttctacaa	cgaccaggcc	gtctactccc	15600
agctcatccg	ccagtttacc	tctctgacct	acgtgttcaa	tcgctttccc	gagaaccaga	15660
ttttggcgcg	cccgccagcc	cccaccatca	ccaccgtcag	tgaaaacgtt	cctgctctca	15720
cagatcacgg	gacgctaccg	ctgcgcaaca	gcacggaggg	agtccagcga	gtgaccatta	15780
ctgacgccag	acgccgcacc	tgcccctacg	tttacaaggc	cctgggcata	gtctcgccgc	15840
gcgtcctatc	gagccgcact	ttttgagcaa	acatgtccat	ccttatatcg	cccagcaata	15900
acacaggctg	gggcctgcgc	ttcccaagca	agatgtttgg	cggggcaaag	aagcgctccg	15960
accaacaccc	agtgcgcgtg	cgcgggcact	accgcgcgcc	ctggggcgcg	cacaaacgcg	16020
gccgcactgg	gcgcaccacc	gtcgatgacg	ccattgacgc	ggtggtggag	gaggcgcgca	16080
actacagccc	cacgccgcca	ccagtgtcca	cagtggacgc	ggccattcag	accgtggtgc	16140
gcggagcccc	gcgttatgct	aaaatgaaga	gacggcggag	gcgcgtagca	cgtcgccacc	16200
gccgccgacc	cggcactgcc	gcccacgcgc	cggcggcggc	cctgcttaac	cgcgcacgtc	16260
gcaccggccg	acgggcggcc	atgcggggcg	ctcgaaggct	ggccgcgggt	attgtcactg	16320
tgccccccag	gtccaggcga	cgagcggccg	ccgcagcagc	cgcggccatt	agtgctatga	16380
ctcagggtcg	caggggcaac	gtgtactggg	tgcgcgactc	ggttagcggc	ctgcgcgtgc	16440
ccgtgcgcac	ccgccccccg	cgcaactaga	ttgcaagaaa	aaactactta	gactcgtact	16500
gttgatatgta	tccagcggcg	gcggcgcgca	acgaagctat	gtccaagcgc	aaaatcaaag	16560
aagagatgct	ccaggtcac	gcgccggaga	tctatggccc	cccgaagaag	gaagagcagg	16620
attacaagcc	ccgaaagcta	aagcgggtca	aaaagaaaaa	gaaagatgat	gatgatgatg	16680
aacttgacga	cgaggtggaa	ctgctgcacg	caaccgcgcc	cagggcgcg	gtacagtggg	16740
aaggtcgacg	cgtaagacgt	gttttgcgac	ccggcaccac	cgtagttttt	acgcccgggtg	16800
agcgtccac	ccgcacctac	aagcgcgtgt	atgatgaggt	gtacggcgac	gaggacctgc	16860
ttgagcaggc	caacgagcgc	ctcggggagt	ttgcctacgg	aaagcggcat	aaggacatgt	16920
tggcgttgcc	gctggacgag	ggcaacccaa	cacctagcct	aaagcccgtg	acactgcagc	16980
aggtgctgcc	cacgcttgca	ccgtccgaag	aaaagcgcg	cctaaagcgc	gagtctggtg	17040

acttggcacc	caccgtgcag	ctgatggtac	ccaagcgcca	gcgactggaa	gatgtcttgg	17100
aaaaaatgac	cgtggagcct	gggctggagc	ccgaggtccg	cgtgcggcca	atcaagcagg	17160
tggcaccggg	actgggcgtg	cagaccgtgg	acgttcagat	accaccacc	agtagcacta	17220
gtattgccac	tgccacagag	ggcatggaga	cacaaacgtc	cccggttgcc	tcggcgggtg	17280
cagatgccgc	ggtgcaggcg	gccgctgcgg	ccgcgtccaa	aacctctacg	gaggtgcaaa	17340
cggacccgtg	gatgtttcgc	gtttcagccc	cccggcgccc	gcgccgttcc	aggaagtacg	17400
gcaccgccag	cgcactactg	cccgaatatg	ccctacatcc	ttccatcgcg	cctacccccg	17460
gctatcgtgg	ctacacctac	cgccccagaa	gacgagcgac	taccgcgacg	cgaaccacca	17520
ctggaacccg	ccgccgcgct	cgccgtcgcc	agcccgtgct	ggccccgatt	tccgtgcgca	17580
gggtggctcg	cgaaggaggc	aggaccctgg	tgctgccaac	agcgcgctac	caccccagca	17640
tcgtttaaaa	gccggtcttt	gtggttcttg	cagatatggc	cctcacctgc	cgctccggtt	17700
tcccggtgcc	gggattccga	ggaagaatgc	accgtaggag	gggcatggcc	ggccacggcc	17760
tgacgggctg	catgcgtcgt	gcgcaccacc	ggcggcgggc	cgcgtcgcac	cgtcgcatgc	17820
gcggcggtat	cctgcccctc	cttattccac	tgatcgccgc	ggcgattggc	gccgtgcccg	17880
gaattgcatc	cgtggccttg	caggcgacga	gacactgatt	aaaaacaagt	tgcatgtgga	17940
aaaatcaaaa	taaaaagtct	ggagtctcac	gctcgcttgg	tcctgtaact	atthttgtaga	18000
atggaagaca	tcaactttgc	gtctctggcc	ccgcgacacg	gctcgcgccc	gttcatggga	18060
aactggcaag	atatcggcac	cagcaatatg	agcggtgggc	ccttcagctg	gggctcgctg	18120
tggagcggca	ttaaaaattt	cggttccacc	attaagaact	atggcagcaa	ggcctggaac	18180
agcagcacag	gccagatgct	gagggacaag	ttgaaagagc	aaaatttcca	acaaaagggtg	18240
gtagatggcc	tggcctctgg	cattagcggg	gtggtggacc	tggccaacca	ggcagtgcaa	18300
aataagatta	acagtaagct	tgatccccgc	cctcccgtag	aggagcctcc	accggccgtg	18360
gagacagtgt	ctccagaggg	gcgtggcgaa	aagcgtccgc	ggcccagacg	ggaagaaact	18420
ctggtgacgc	aaatagatga	gcctccctcg	tacgaggagg	cactaaagca	aggcctgccc	18480
accacccgtc	ccatcgcgcc	catggctacc	ggagtgcctg	gccagcacac	acctgtaacg	18540
ctggacctgc	ctccccccgc	tgacacccag	cagaaacctg	tgctgccagg	gccgtccgcc	18600
gttggtgtaa	ccgcgccctag	ccgcgcgtcc	ctgcgcctg	ccgccagcgg	tccgcgatcg	18660
atgcggcccc	tagccagtgg	caactggcaa	agcacactga	acagcatcgt	gggtctgggg	18720
gtgcaatccc	tgaagcgccg	acgatgcttc	taaatagcta	acgtgtcgta	tgtgtcatgt	18780
atgcgtccat	gtcgccgcca	gaggagctgc	tgagccgccc	tgcgcccgtc	ttccaagatg	18840
gctacccctt	cgatgatgcc	gcagtgggtc	tacatgcaca	tctcgggcca	ggacgcctcg	18900

gagtacctga	gccccgggct	ggtgcagttt	gcccgcgcca	ccgagacgta	cttcagcctg	18960
aataacaagt	ttagaaaccc	cacggtggca	cctacgcacg	acgtaaccac	agaccgggtcc	19020
cagcgtttga	cgctgcgggt	catccctgtg	gaccgcgagg	ataccgcgta	ctcgtacaaa	19080
gcgcggttca	ccctggctgt	gggtgacaac	cgtgtgcttg	atatggcttc	cacgtacttt	19140
gacatccgcg	gcgtgctgga	cagggggcct	acttttaagc	cctactccgg	cactgcctac	19200
aacgctctag	ctcccaaggg	cgctcctaac	tcctgtgagt	gggaacaaac	cgaagatagc	19260
ggccgggacg	ttgccgagga	tgaagaagag	gaagatgaag	atgaagaaga	ggaagaagaa	19320
gagcaaaacg	ctcgagatca	ggctactaag	aaaacacatg	tctatgcca	ggctcctttg	19380
tctggagaaa	caattacaaa	aagcgggcta	caaataggat	cagacaatgc	agaaacacaa	19440
gctaaacctg	tatacgcaga	tccttcctat	caaccagaac	ctcaaattgg	cgaatctcag	19500
tggaacgaag	ctgatgctaa	tgccgcagga	gggagagtgc	ttaaaaaac	aactcccatg	19560
aaacctatgct	atggatctta	tgccaggcct	acaaatcctt	ttggtggtca	atccgttctg	19620
gttccggatg	aaaaaggggt	gcctcttcca	aagggtgact	tgcaattctt	ctcaaatact	19680
acctctttga	acgaccggca	aggcaatgct	actaaaccaa	aagtggtttt	gtacagtgaa	19740
gatgtaaata	tggaaacccc	agacacacat	ctgtcttaca	aacctggaaa	aggtgatgaa	19800
aattctaaag	ctatgttggg	tcaacaatct	atgccaaaca	gaccaatta	cattgctttc	19860
agggacaatt	ttattggcct	aatgtattat	aacagcactg	gcaacatggg	tgttcttgct	19920
ggtcaggcat	cgcagctaaa	tgccgtggta	gatttgcaag	acagaaacac	agagctgtcc	19980
tatcaactct	tgcttgattc	cataggtgat	agaaccagat	atctttctat	gtggaatcag	20040
gctgtagaca	gctatgatcc	agatgttaga	atcattgaaa	accatggaac	tgaggatgaa	20100
ttgccaaatt	attgttttcc	tcttgggggt	attggggtaa	ctgacaccta	tcaagctatt	20160
aaggctaattg	gcaatggctc	aggcgataat	ggagatacta	catggacaaa	agatgaaact	20220
tttgcaacac	gtaatgaaat	aggagtgggt	aacaactttg	ccatggaaat	taacctaaat	20280
gccaacctat	ggagaaatth	cctttactcc	aatattgcgc	tgtacctgcc	agacaagcta	20340
aaatacaacc	ccaccaatgt	ggaaatatct	gacaacccca	acacctacga	ctacatgaac	20400
aagcgagtgg	tggctcccgg	gcttgtagac	tgctacatta	accttggggc	gcgctggtct	20460
ctggactaca	tggacaacgt	taatcccttt	aaccaccacc	gcaatgcggg	cctccgttat	20520
cgctccatgt	tgttgggaaa	cggccgctac	gtgccctttc	acattcaggt	gccccaaaag	20580
ttttttgcca	ttaaaaacct	cctcctcctg	ccaggctcat	atacatatga	atggaacttc	20640
aggaaggatg	ttaacatggt	tctgcagagc	tctctgggaa	acgatcttag	agttgacggg	20700
gctagcatta	agtttgacag	catttgtctt	tacgccacct	tcttcccat	ggcccacaac	20760
acggcctcca	cgctggaagc	catgctcaga	aatgacacca	acgaccagtc	ctttaatgac	20820

tacctttccg	ccgccaacat	gctatacccc	ataccgcgca	acgccaccaa	cgtgcccac	20880
tccatcccat	cgcgcaactg	ggcagcattt	cgcgggttggg	ccttcacacg	cttgaagaca	20940
aaggaaaccc	cttccctggg	atcaggctac	gacccttact	acacctactc	tggctccata	21000
ccataccttg	acggaacctt	ctatcttaat	cacaccttta	agaagggtggc	cattaccttt	21060
gactcttctg	ttagctggcc	gggcaacgac	cgcctgctta	ctcccaatga	gtttgagatt	21120
aaacgctcag	ttgacgggga	gggctacaac	gtagctcagt	gcaacatgac	caaggactgg	21180
ttcctgggtg	agatgttggc	caactacaat	attggctacc	agggcttcta	cattccagaa	21240
agctacaagg	accgcatgta	ctcgttcttc	agaaacttcc	agcccatgag	ccggcaagtg	21300
gttgacgata	ctaaatacaa	ggagtatcag	caggttggaa	ttcttcacca	gcataacaac	21360
tcaggattcg	taggctacct	cgtctccacc	atgcgcgagg	gacaggctta	ccccgccaac	21420
gtgccctacc	cactaatagg	caaaaccgcg	gttgacagta	ttaccagaa	aaagtttctt	21480
tgcgatcgca	ccctttggcg	catcccatc	tccagtaact	ttatgtccat	gggcgcactc	21540
acagacctgg	gccaaaacct	tctctacgcc	aactccgccc	acgcgctaga	catgactttt	21600
gaggtggatc	ccatggacga	gccaccctt	ctttatgttt	tgtttgaagt	ctttgacgtg	21660
gtccgtgtgc	accagccgca	ccgcggcgtc	atcgagaccg	tgtacctgcg	cacgcccttc	21720
tcggccggca	acgccacaac	ataaaagaag	caagcaacat	caacaacagc	tgccgccatg	21780
ggctccagtg	agcaggaact	gaaagccatt	gtcaaagatc	ttgggttgtgg	gccatatttt	21840
ttgggcacct	atgacaagcg	ctttccaggc	tttgtttctc	cacacaagct	cgcctgcgcc	21900
atagtcaata	cggccggtcg	cgagactggg	ggcgtacact	ggatggcctt	tgcctggaac	21960
ccgcgctcaa	aaacatgcta	cctctttgag	ccctttggct	tttctgacca	acgactcaag	22020
caggtttacc	agtttgagta	cgagtcactc	ctgcgcgta	gcgccattgc	ttcttcccc	22080
gaccgctgta	taacgctgga	aaagtccacc	caaagcgtgc	aggggcccaa	ctcggccgcc	22140
tgtggactat	tctgctgcat	gtttctccac	gcctttgcca	actggcccca	aactcccatg	22200
gatcacaacc	ccaccatgaa	ccttattacc	ggggtaccca	actccatgct	taacagtccc	22260
caggtagacg	ccaccctgcg	tgcgaaccag	gaacagctct	acagcttcct	ggagcgccac	22320
tgcacctact	tccgcagcca	cagtgcgcag	attaggagcg	ccacttcttt	ttgtcacttg	22380
aaaaacatgt	aaaaataatg	tactaggaga	cactttcaat	aaaggcaa	gtttttat	22440
gtacactctc	gggtgattat	ttacccccca	cccttgccgt	ctgcgcggtt	taaaaatcaa	22500
aggggttctg	ccgcgcatcg	ctatgcgcca	ctggcaggga	cacgttgcca	tactggtg	22560
tagtgctcca	cttaaactca	ggcacaacca	tccgcggcag	ctcgggtgaag	ttttcactcc	22620
acaggctgcg	caccatcacc	aacgcgttta	gcaggtcggg	cgccgatatc	ttgaagtcgc	22680

agttggggcc	tccgccctgc	gcgcgcgagt	tgcgatacac	agggttgcag	caactggaaca	22740
ctatcagcgc	cgggttggtgc	acgctggcca	gcacgctctt	gtcggagatc	agatccgcgt	22800
ccaggtcctc	cgcgttgctc	agggcgaacg	gagtcaactt	tggtagctgc	cttcccaaaa	22860
aggggtgcatg	cccaggcttt	gagttgcact	cgcaccgtag	tggcatcaga	aggtgaccgt	22920
gcccggctctg	ggcgttagga	tacagcgctt	gcatgaaagc	cttgatctgc	ttaaaagcca	22980
cctgagcctt	tgcgccttca	gagaagaaca	tgccgcaaga	cttgccggaa	aactgattgg	23040
ccggacaggc	cgcgtcatgc	acgcagcacc	ttgcgtcggg	gttggagatc	tgcaccacat	23100
ttcggcccca	ccggttcttc	acgatcttgg	ccttgctaga	ctgctccttc	agcgcgcgct	23160
gcccgttttc	gctcgtcaca	tccatttcaa	tcacgtgctc	cttatttatc	ataatgctcc	23220
cgtgtagaca	cttaagctcg	ccttcgatct	cagcgcagcg	gtgcagccac	aacgcgcagc	23280
ccgtgggctc	gtggtgcttg	taggttacct	ctgcaaacga	ctgcaggtac	gcctgcagga	23340
atcgccccat	catcgtcaca	aaggctcttg	tgctgggtgaa	ggtcagctgc	aaccgcgggt	23400
gctcctcggt	tagccaggtc	ttgcatacgg	ccgccagagc	ttccacttgg	tcaggcagta	23460
gcttgaagtt	tgcctttaga	tcgttatcca	cgtgggtactt	gtccatcaac	gcgcgcgcag	23520
cctccatgcc	cttctcccac	gcagacacga	tcggcaggct	cagcggggtt	atcacctgct	23580
tttcaactttc	cgtctcactg	gactcttctt	tttctctctg	cgtccgcata	ccccgcgcca	23640
ctgggtcgtc	ttcattcagc	cgcgcgacgg	tgcgcttacc	tcccttgccg	tgcttgatta	23700
gcaccggtgg	gttgctgaaa	cccaccattt	gtagcgccac	atcttctctt	tcttcctcgc	23760
tgtccacgat	cacctctggg	gatggcgggc	gctcgggctt	gggagagggg	cgttctcttt	23820
tctttttgga	cgcaatggcc	aaatccgcgg	tcgaggtcga	tggccgcggg	ctgggtgtgc	23880
gcggcaccag	cgcactctgt	gacgagctct	cttcgtcctc	ggactcgaga	cgcgcctca	23940
gccgcttttt	tgggggcgcg	cggggaggcg	gcggcgacgg	cgcgggggac	gacacgtcct	24000
ccatggttgg	tggacgtcgc	gccgcaccgc	gtccgcgctc	gggggtgggt	tcgcgctgct	24060
cctcttcccg	actggccatt	tccttctcct	ataggcagaa	aaagatcatg	gagtcagtcg	24120
agaaggagga	cagcctaacc	gccccctttg	agttcgccac	caccgcctcc	accgatgccg	24180
ccaacgcgcc	taccaccttc	cccgtcgagg	cacccccgct	tgaggaggag	gaagtgatta	24240
tcgagcagga	cccaggtttt	gtaagcgaag	acgacgagga	tcgctcagta	ccaacagagg	24300
ataaaaagca	agaccaggac	gacgcagagg	caaacgagga	acaagtcggg	cggggggacc	24360
aaaggcatgg	cgactacctt	gatgtgggag	acgacgtgct	gttgaagcat	ctgcagcgcc	24420
agtgcgccat	tatctgcgac	gcgttgcaag	agcgcagcga	tgtgcccctc	gccatagcgg	24480
atgtcagcct	tgcctacgaa	cgccacctgt	tctcaccgcg	cgtaccccc	aaacgccaaag	24540
aaaacggcac	atgcgagccc	aaccgcggcc	tcaacttcta	ccccgtattt	gccgtgccag	24600

agggtgcttgc	cacctatcac	atctttttcc	aaaactgcaa	gataccccta	tcctgccgtg	24660
ccaaccgcag	ccgagcggac	aagcagctgg	ccttgcgcca	gggcgctgtc	atacctgata	24720
tcgcctcgct	cgacgaagtg	ccaaaaatct	ttgaggggtct	tggacgcgac	gagaaacgcg	24780
cggcaaacgc	tctgcaacaa	gaaaacagcg	aaaatgaaag	tactgtgga	gtgctggtgg	24840
aacttgaggg	tgacaacgcg	cgcctagccg	tgctgaaacg	cagcatcgag	gtcaccact	24900
ttgcctaccc	ggcacttaac	ctacccccca	aggttatgag	cacagtcag	agcgagctga	24960
tcgtgcgccc	tgacgacccc	ctggagaggg	atgcaaacct	gcaagaacaa	accgaggagg	25020
gcctacccgc	agttggcgat	gagcagctgg	cgcgctggct	tgagacgcgc	gagcctgccg	25080
acttgaggga	gcgacgcaag	ctaatgatgg	ccgcagtgct	tgttaccgtg	gagcttgagt	25140
gcatgcagcg	gttctttgct	gacccggaga	tgacgcgcaa	gctagaggaa	acgttgcaact	25200
acacctttcg	ccagggctac	gtgcgccagg	cctgcaaaat	ttccaacgtg	gagctctgca	25260
acctggtctc	ctaccttgga	atcttgcaag	aaaaccgcct	cgggcaaac	gtgcttcatt	25320
ccacgctcaa	gggcgaggcg	cgcgcgcact	acgtccgcga	ctgcgtttac	ttatttctgt	25380
gctacacctg	gcaaacggcc	atgggcgtgt	ggcagcaatg	cctggaggag	cgcaacctaa	25440
aggagctgca	gaagctgcta	aagcaaaact	tgaaggacct	atggacggcc	ttcaacgagc	25500
gctccgtggc	cgcgcacctg	gcggaacatta	tcttccccga	acgcctgctt	aaaaccctgc	25560
aacaggggtct	gccagacttc	accagtcaaa	gcatgttgca	aaactttagg	aactttatcc	25620
tagagcgttc	aggaattctg	cccgccacct	gctgtgcgct	tcctagcgac	tttgtgccca	25680
ttaagtaccg	tgaatgccct	ccgcgcgttt	ggggtcactg	ctaccttctg	cagctagcca	25740
actaccttgc	ctaccactcc	gacatcatgg	aagacgtgag	cggtgacggc	ctactggagt	25800
gtcactgtcg	ctgcaacctc	tgacccccgc	accgctccct	ggtctgcaat	tcgcaactgc	25860
ttagcgaaag	tcaaattatc	ggtacctttg	agctgcaggg	tccctcgcct	gacgaaaagt	25920
ccgcggctcc	gggggttga	ctcactccgg	ggctgtggac	gtcggcttac	cttcgcaaat	25980
ttgtacctga	ggactaccac	gcccacgaga	ttaggttcta	cgaagaccaa	tcccgcccgc	26040
caaatgcgga	gcttaccgcc	tgcgtcatta	cccagggcca	catccttggc	caattgcaag	26100
ccatcaacaa	agcccgccaa	gagtttctgc	tacgaaaggg	acgggggggtt	tacctggacc	26160
cccagtcgag	cgaggagctc	aacccaatcc	ccccgcgcgc	gcagccctat	cagcagccgc	26220
gggcccttgc	ttcccaggat	ggcaccacaa	aagaagctgc	agctgccgcc	gccgccaccc	26280
acggacgagg	aggaatactg	ggacagtcag	gcagaggagg	ttttggacga	ggaggaggag	26340
atgatggaag	actgggacag	cctagacgaa	gcttccgagg	ccgaagaggt	gtcagacgaa	26400
acaccgtcac	cctcggtcgc	attcccctcg	ccggcgcccc	agaaattggc	aaccgttccc	26460

agcatcgcta	caacctccgc	tcctcaggcg	ccgccggcac	tgcctgttcg	ccgacccaac	26520
cgtagatggg	acaccactgg	aaccagggcc	ggtaagtcta	agcagccgcc	gccgttagcc	26580
caagagcaac	aacagcgcca	aggctaccgc	tcgtggcgcg	ggcacaagaa	cgccatagtt	26640
gcttgcttgc	aagactgtgg	gggcaacatc	tccttcgccc	gccgctttct	tctctaccat	26700
cacggcgtgg	ccttcccccg	taacatcctg	cattactacc	gtcatctcta	cagcccctac	26760
tgcaccggcg	gcagcggcag	cggcagcaac	agcagcggtc	acacagaagc	aaaggcgacc	26820
ggatagcaag	actctgacaa	agcccaagaa	atccacagcg	gcggcagcag	caggaggagg	26880
agcgctgcgt	ctggcgccca	acgaaccctg	atcgaccctg	gagcttagaa	ataggatttt	26940
tcccactctg	tatgctatat	ttcaacaaag	cagggggcaa	gaacaagagc	tgaaaataaa	27000
aaacaggtct	ctgcgctccc	tcaccgcgag	ctgcctgtat	cacaaaagcg	aagatcagct	27060
tcggcgcacg	ctggaagacg	cggaggctct	cttcagcaaa	tactgcgcgc	tgactcttaa	27120
ggactagttt	cgcgcccttt	ctcaaattta	agcgcgaaaa	ctacgtcatc	tccagcggcc	27180
acaccggcg	ccagcacctg	tcgtcagcgc	cattatgagc	aaggaaattc	ccacgcctta	27240
catgtggagt	taccagccac	aaatgggact	tgcggctgga	gctgcccaag	actactcaac	27300
ccgaataaac	tacatgagcg	cgggacccca	catgatatcc	cgggtcaacg	gaatccgcgc	27360
ccaccgaaac	cgaattctcc	tcgaacaggc	ggctattacc	accacacctc	gtaataacct	27420
taatccccgt	agttggcccc	ctgccctggg	gtaccaggaa	agtcccgcgc	ccaccactgt	27480
ggtacttccc	agagacgccc	aggccgaagt	tcagatgact	aactcagggg	cgcagcttgc	27540
gggcggcttt	cgtcacaggg	tgcggctgcc	cgggcagggt	ataactcacc	tgaaaatcag	27600
agggcgaggt	attcagctca	acgacgagtc	ggtgagctcc	tctcttggtc	tccgtccgga	27660
cgggacattt	cagatcggcg	gcgctggccg	ctcttcattt	acgccccgtc	aggcgatcct	27720
aactctgcag	acctcgtcct	cggagccgcg	ctccggaggc	attggaactc	tacaatttat	27780
tgaggagttc	gtgccttcgg	tttacttcaa	ccccttttct	ggacctcccg	gccactaccc	27840
ggaccagttt	attcccaact	ttgacgcggg	gaaagactcg	gcggacggct	acgactgaat	27900
gaccagtgga	gaggcagagc	gactgcgcct	gacacacctc	gaccactgcc	gccgccacaa	27960
gtgctttgcc	cgcggctccg	gtgagttttg	ttactttgaa	ttgcccgaa	agcatatcga	28020
gggcccggcg	cacggcgctc	ggctcaccac	ccaggtagag	cttacacgta	gcctgattcg	28080
ggagtttacc	aagcgccccc	tgctagtgga	gcgggagcgg	ggtccctgtg	ttctgaccgt	28140
ggtttgcaac	tgtcctaacc	ctggattaca	tcaagatctt	tgttgctcatc	tctgtgctga	28200
gtataataaa	tacagaaatt	agaatctact	ggggctcctg	tcgccatcct	gtgaacgcca	28260
ccgtttttac	ccacccaaag	cagaccaaag	caaacctcac	ctccggtttg	cacaagcggg	28320
ccaataagta	ccttacctgg	tactttaacg	gctcttcatt	tgtaatattac	aacagtttcc	28380

agcgagacga	agtaagtttg	ccacacaacc	ttctcggctt	caactacacc	gtcaagaaaa	28440
acaccaccac	caccaccctc	ctcacctgcc	gggaacgtac	gagtgcgtca	ccggttgctg	28500
cgccacacac	tacagcctga	gcgtaaccag	acattactcc	cattttttcca	aaacaggagg	28560
tgagctcaac	tcccggaaact	caggtcaaaa	aagcattttg	cggggtgctg	ggatttttta	28620
attaagtata	tgagcaattc	aagtaactct	acaagcttgt	ctaatttttc	tggaattggg	28680
gtcgggggtta	tccttactct	tgtaattctg	tttattctta	tactagcact	tctgtgcctt	28740
aggggttgccg	cctgctgcac	gcacgtttgt	acctattgtc	agctttttta	acgctggggg	28800
caacatccaa	gatgaggtag	atgatttttag	gcttgctcgc	ccttgccggca	gtctgcagcg	28860
ctgccaaaaa	ggttgagttt	aaggaaccag	cttgcaatgt	tacattttaa	tcagaagcta	28920
atgaatgcac	tactcttata	aaatgcacca	cagaacatga	aaagcttatt	attcgccaca	28980
aagacaaaat	tggcaagtat	gctgtatatg	ctatttggca	gccaggtagc	actaacgact	29040
ataatgtcac	agtcttccaa	ggtgaaaatc	gtaaaacttt	tatgtataaa	tttccatttt	29100
atgaaatgtg	cgatattacc	atgtacatga	gcaaacagta	caagttgtgg	ccccacaaa	29160
agtgtttaga	gaacactggc	accttttggt	ccaccgctct	gcttattaca	gcgcttgctt	29220
tggtatgtac	cttactttat	ctcaaataca	aaagcagacg	cagttttatt	gatgaaaaga	29280
aaatgccttg	attttccgct	tgcttgtatt	cccctggaca	atttactcta	tgtgggatat	29340
gctccaggcg	ggcaagatta	taccacacaac	cttcaaataca	aactttcctg	gacgttagcg	29400
cctgattttct	gccagcgcct	gcactgcaaa	tttgatcaaa	cccagcttca	gcttgccctgc	29460
tccagagatg	accggtctca	ccatcgcgcc	cacaacggac	tatcgcaaca	ccactgctac	29520
cggactaaca	tctgccctaa	atttacccca	agttcatgcc	tttgtcaatg	actgggagag	29580
cttggacatg	tggtgggttt	ccatagcgct	tatgtttggt	tgccttatta	ttatgtggct	29640
tatttgttgc	ctaaagcgca	gacgcgccag	accccccatc	tataggccta	tcattgtgct	29700
caaccacac	aatgaaaaaa	ttcatagatt	ggacgggtctg	aaaccatggt	ctcttctttt	29760
acagtatgat	taaatgagac	atgattcctc	gagttcttat	attattgacc	cttgttgccg	29820
ttttctgtgc	gtgctctaca	ttggccggcg	tcgctcacat	cgaagtagat	tgcatccac	29880
ctttcacagt	ttacctgctt	tacggatttg	tcacccttat	cctcatctgc	agcctcgtca	29940
ctgtagtcat	cgccttcatt	cagttcattg	actgggtttg	tgtgcgcatt	gcgtacctca	30000
ggcaccatcc	gcaatacaga	gacaggacta	tagctgatct	tctcagaatt	ctttaattat	30060
gaaacggagt	gtcatttttg	ttttgctgat	tttttgccgc	ctacctgtgc	tttgctccca	30120
aacctcagcg	cctcccaaaa	gacatatttc	ctgcagattc	actcaaatat	ggaacattcc	30180
cagctgctac	aacaaacaga	gcgatttgtc	agaagcctgg	ttatacgcca	tcattctctgt	30240

catggttttt	tgcagtagca	tttttgccct	agccatatat	ccataccttg	acattggctg	30300
gaatgccata	gatgccatga	accaccctac	tttcccagtg	cccgtgtgca	taccactgca	30360
acaggttatt	gccccaatca	atcagcctcg	cccccttct	cccacccccca	ctgagattag	30420
ctactttaat	ttgacagggtg	gagatgactg	aatctctaga	tctagaattg	gatggaatta	30480
acaccgaaca	gcgcctacta	gaaaggcgca	aggcgggcgtc	cgagcgagaa	cgcctaaaac	30540
aagaagttga	agacatgggtt	aacctacacc	agtgtaaaag	aggatatcttt	tgtgtgggtca	30600
agcaggccaa	acttacctac	gaaaaaacca	ctaccggcaa	ccgcctcagc	tacaagctac	30660
ccaccagcg	ccaaaaactg	gtgcttatgg	tgggagaaaa	acctatcacc	gtcaccagc	30720
actcggcaga	aacagagggc	tgcctgcact	tcccctatca	gggtccagag	gacctctgca	30780
ctcttattaa	aacatgtgt	ggtattagag	atcttattcc	attcaactaa	cataaacaca	30840
caataaatta	cttacttaaa	atcagtcagc	aaatctttgt	ccagcttatt	cagcatcacc	30900
tcctttcctt	cctcccaact	ctggtatctc	agccgccttt	tagctgcaaa	ctttctccaa	30960
agtttaaatg	ggatgtcaaa	ttcctcatgt	tcttgtccct	ccgcacccac	tatcttcata	31020
ttgttgacga	tgaaacgcgc	cagaccgtct	gaagacacct	tcaaccccgt	gtatccatat	31080
gacacagaaa	ccgggcctcc	aactgtgccc	tttcttacct	ctccatttgt	ttcacccaat	31140
ggtttccaag	aaagtccccc	tggagttctc	tctctacgcg	tctccgaacc	tttggacacc	31200
tcccacggca	tgttgcgct	taaaatgggc	agcgggtctta	ccctagacaa	ggccggaaac	31260
ctcacctccc	aaaatgtaac	cactgttact	cagccactta	aaaaaacaaa	gtcaaacata	31320
agtttgacga	cctccgcacc	acttacaatt	acctcaggcg	ccctaacagt	ggcaaccacc	31380
gctcctctga	tagttactag	cggcgctctt	agcgtacagt	cacaagcccc	actgaccgtg	31440
caagactcca	aactaagcat	tgtactaaa	gggccattta	cagtgtcaga	tggaaagcta	31500
gccctgcaaa	catcagcccc	cctctctggc	agtgcagcgc	acacccttac	tgttaactgca	31560
tcacccccgc	taactactgc	cacgggtagc	ttggggcatta	acatggaaga	tcctatattat	31620
gtaaataatg	gaaaaatagg	aattaaaata	agcgggtcctt	tgcaagtagc	acaaaactcc	31680
gatacactaa	cagtagttac	tggaccaggt	gtcaccgttg	aacaaaactc	ccttagaacc	31740
aaagttgcag	gagctattgg	ttatgattca	tcaaacaaca	tggaaattaa	aacggggcgg	31800
ggcatgcgta	taaataacaa	cttggttaatt	ctagatgtgg	attaccatt	tgatgctcaa	31860
acaaaactac	gtcttaaaact	ggggcagggg	cccctgtata	ttaatgcatc	tcataacttg	31920
gacataaaact	ataacagagg	cctatacctt	tttaatgcat	caaacaatac	taaaaaactg	31980
gaagtttagca	taaaaaaatc	cagtggacta	aactttgata	atactgccat	agctataaat	32040
gcaggaaaagg	gtctggaggtt	tgatacaaac	acatctgagt	ctccagatat	caacccaata	32100
aaaactaaaa	ttggctctgg	cattgattac	aatgaaaacg	gtgccatgat	tactaaactt	32160

ggagcggggtt taagctttga caactcaggg gccattacaa taggaaacaa aaatgatgac	32220
aaacttacctt tgtggacaac ccagaccca tctcctaact gcagaattca ttcagataat	32280
gactgcaaatt ttacttttgt tcttacaaaa tgtggggagtc aagtactagc tactgtagct	32340
gctttggctg tatctggaga tctttcatcc atgacaggca ccgttgcaag tgttagtata	32400
ttccttagat ttgaccaaaa cgggtgttcta atggagaact cctcacttaa aaaacattac	32460
tggaacttta gaaatgggaa ctcaactaat gcaaatccat acacaaatgc agttggattt	32520
atgcctaacc ttctagccta tccaaaaacc caaagtcaaa ctgctaaaaa taacattgtc	32580
agtcaagttt acttgcctgg tgataaaact aaacctatga tacttaccat tacacttaat	32640
ggcactagtg aatccacaga aactagcgag gtaagcactt actctatgtc ttttacatgg	32700
tcctgggaaa gtggaaaata caccactgaa acttttgcta ccaactctta caccttctcc	32760
tacattgccc aggaataaag aatcgtgaac ctggtgcatg ttatgtttca acgtgtttat	32820
ttttcaattg cagaaaattt caagtcattt ttcattcagt agtatagccc caccaccaca	32880
tagcttatat tgatcacctg accttaatac aactcacaga accctagtat tcaacctgcc	32940
acctccctcc caacacacag agtacacagt cctttctccc cggctggcct taaaaagcat	33000
catatcatgg gtaacagaca tattcttagg tgttatattc cacacgggtt cctgtcgagc	33060
caaacgctca tcagtgatat taataaaact cccgggcagc tcgcttaagt tcatgtcgct	33120
gtccagctgc tgagccacag gctgctgtcc aacttgccgt tgctcaacgg gcggcgaagg	33180
ggaagtccac gcctacatgg gggtagagtc ataatcgtgc atcaggatag ggcggtggtg	33240
ctgcagcagc gcgcgaataa actgctgccg ccgcccgtcc gtccctgcagg aatacaacat	33300
ggcagtggtc tcctcagcga tgattcgcac cgcgcgcagc atgagacgcc ttgtcctccg	33360
ggcacagcag cgcaccctga tctcacttaa atcagcacag taactgcagc acagcaccac	33420
aatattgttc aaaatccac agtgcaaggc gctgtatcca aagctcatgg cggggaccac	33480
agaacccacg tggccatcat accacaagcg caggtagatt aagtggcgac ccctcataaa	33540
cacgctggac ataaacatta cctcttttgg catgttgtaa ttcaccacct cccggtacca	33600
tataaacctc tgattaaaca tggcgccatc caccaccatc ctaaaccagc tggccaaaac	33660
ctgcccgcg gctatgcact gcagggaacc gggactggaa caatgacagt ggagagccca	33720
ggactcgtaa ccatggatca tcatgctcgt catgatatca atgttggcac aacacaggca	33780
cacgtgcata cacttcctca ggattacaag ctccctccgc gtcagaacca tatcccagg	33840
aacaacccat tcctgaatca gcgtaaatcc cacactgcag ggaagacctc gcacgtaact	33900
cacgttgtgc attgtcaaag tgttacattc gggcagcagc ggatgatcct ccagtatggt	33960
agcgcgggtc tctgtctcaa aaggaggtag gcgatcccta ctgtacggag tgcgccgaga	34020

caaccgagat	cgtgttggtc	gtagtgtcat	gccaaatgga	acgccggacg	tagtcatatt	34080
tcctgaagca	aaaccaggtg	cgggcgtgac	aaacagatct	gcgtctccgg	tctcgtcgct	34140
tagctcgctc	tgtgtagtag	ttgtagtata	tccactctct	caaagcatcc	aggcgccccc	34200
tggcttcggg	ttctatgtaa	actccttcat	gcgcgctgc	cctgataaca	tccaccaccg	34260
cagaataagc	cacaccacgc	caacctacac	attcgttctg	cgagtcacac	acgggaggag	34320
cgggaagagc	tggaagaacc	atgttttttt	tttttttatt	caaaagatt	atccaaaacc	34380
tcaaaatgaa	gatctattaa	gtgaacgcgc	ccccctccgg	tggcgtggtc	aaactctaca	34440
gccaaagaac	agataatggc	atttgtaaga	tgttgacaaa	tggcttccaa	aaggcaaact	34500
gccctcacgt	ccaagtggac	gtaaaggcta	aacccttcag	ggtgaatctc	ctctataaac	34560
attccagcac	cttcaaccat	gccccaaata	ttttcatctc	gccaccttat	caatatgtct	34620
ctaagcaaatt	cccgaatatt	aagtcgggcc	attgtaaaaa	tctgctccag	agcgccctcc	34680
accttcagcc	tcaagcagcg	aatcatgatt	gcaaaaattc	aggttcctca	cagacctgta	34740
taagattcaa	aagcgggaaca	ttaacaaaaa	taccgcgatc	ccgtaggtcc	cttcgcaggg	34800
ccagctgaac	ataatcgtgc	aggtctgcac	ggaccagcgc	ggccacttcc	ccgccaggaa	34860
ccatgacaaa	agaaccacac	ctgattatga	cacgcatact	cggagctatg	ctaaccagcg	34920
tagccccctat	gtaagcttgt	tgcattggcg	gcgatataaa	atgcaagggtg	ctgctcaaaa	34980
aatcaggcaa	agcctcgcgc	aaaaaagcaa	gcacatcgta	gtcatgctca	tgcagataaa	35040
ggcaggtaag	ttccggaacc	accacagaaa	aagacaccat	ttttctctca	aacatgtctg	35100
cgggttcctg	cattaaacac	aaaataaaat	aacaaaaaaa	aacatttaaa	cattagaagc	35160
ctgtcttaca	acaggaaaaa	caacccttat	aagcataaga	cggactacgg	ccatgccggc	35220
gtgaccgtaa	aaaaactggt	caccgtgatt	aaaaagcacc	accgacagtt	cctcgggtcat	35280
gtccggagtc	ataatgtaag	actcggtaaa	cacatcaggt	tggttaacat	cggtcagtgc	35340
taaaaagcga	ccgaaatagc	ccgggggaat	acatacccg	aggcgtagag	acaacattac	35400
agcccccata	ggaggtataa	caaaattaat	aggagagaaa	aacacataaa	cacctgaaaa	35460
accctcctgc	ctaggcaaaa	tagcacctc	ccgctccaga	acaacataca	gcgcttcac	35520
agcggcagcc	ataacagtca	gccttaccag	taaaaaaacc	tattaaaaaa	caccactcga	35580
cacggcacca	gctcaatcag	tcacagtgt	aaaagggcca	agtacagagc	gagtatatat	35640
aggactaaaa	aatgacgtaa	cgggttaaagt	ccacaaaaaa	caccagaaa	accgcacgcg	35700
aacctacgcc	cagaaacgaa	agccaaaaaa	cccacaactt	cctcaaactc	tcacttccgt	35760
tttcccacga	tacgtcactt	cccatTTTTAA	aaaaactaca	attcccaata	catgcaagtt	35820
actccgccct	aaaacctacg	tcacccggcc	cgttcccacg	ccccgcgcca	cgtcaciaaac	35880
tccacccccct	cattatcata	ttggcttcaa	tccaaaataa	ggtatattat	gatgatg	35937

5	<210> 4		
	<211> 8342		
	<212> ADN		
	<213> Secuencia artificial		
10	<220>		
	<223> Región del gen auxiliar		
	<400> 4		
	ccgtgcaaaa ggagagcctg taagcgggca ctcttccgtg gtctggtgga taaattcgca	60	
	agggatatcat ggcggacgac cgggggttcga gccccgtatc cggccgtccg ccgtgatcca	120	
	tgcggttacc gcccgcggtg cgaaccaggg tgtgcgacgt cagacaacgg gggagtgctc	180	
	cttttggttt ccttccaggc gcggcgggtg ctgcgctagc ttttttggcc actggccgcg	240	
	cgcagcgtaa gcggttaggc tggaaagcga aagcattaag tggtcgcctc cctgtagccg	300	
	gagggttatt ttccaagggg tgagtcgcgg gacccccggg tcgagtcctc gaccggccgg	360	
	actgcggcga acgggggttt gcctccccgt catgcaagac cccgcttgca aattcctccg	420	
	gaaacaggga cgagcccctt ttttgctttt aaaggcaaag gcttttatct gtacactctc	480	
	gggtgattat ttacccccac ccttgccgtc tgcgcggttt aaaaatcaaa ggggttctgc	540	
	cgcgcatcgc tatgcgccac tggcagggac acgttgcgat actggtgttt agtgctccac	600	
	ttaaactcag gcacaaccat ccgcggcgac tcggtgaagt tttcactcca caggtgcgc	660	
	accatcacca acgcgtttag caggtcgggc gccgatctct tgaagtcgca gttggggcct	720	
	ccgccctgcg cgcgcgagtt gcgatacaca gggttgcagc actggaacac tatcagcgcc	780	
	gggtggtgca cgctggccag cacgctcttg tcggagatca gatccgcgtc caggtcctcc	840	
	gcgttgctca gggcgaacgg agtcaacttt ggtagctgcc ttcccaaaaa gggcgcgtgc	900	
	ccaggctttg agttgactc gcaccgtagt ggcacaaaa ggtgaccgtg cccggtctgg	960	
	gcgttaggat acagcgcctg cataaaagcc ttgatctgct taaaagccac ctgagccttt	1020	
	gcgccttcag agaagaacat gccgcaagac ttgccggaaa actgattggc cggacaggcc	1080	
	gcgtcgtgca cgcagcacct tgcgtcgggtg ttggagatct gcaccacatt tcggccccac	1140	
	cggttcttca cgatcttggc cttgctagac tgctccttca gcgcgcgtg cccgttttcg	1200	
	ctcgtcacat ccatttcaat cacgtgctcc ttatttatca taatgcttcc gtgtagacac	1260	
	ttaagctcgc cttcgatctc agcgcagcgg tgcagccaca acgcgcagcc cgtgggctcg	1320	
	tgatgcttgt aggtcacctc tgcaaacgac tgcaggtagc cctgcaggaa tcgccccatc	1380	
	atcgtcacia aggtcttggt gctgggtgaag gtcagctgca acccgcggtg ctccctcgttc	1440	
	agccaggctc tgcatacggc cgccagagct tccacttggt caggcagtag tttgaagtcc	1500	
	gccttttagat cgttatccac gtggtacttg tccatcagcg cgcgcgcagc ctccatgccc	1560	

ttctcccacg	cagacacgat	cggcacactc	agcgggttca	tcaccgtaat	ttcactttcc	1620
gcttcgctgg	gctcttcctc	ttcctcttgc	gtccgcatac	cacgcgccac	tgggtcgtct	1680
tcattcagcc	gccgcactgt	gcgcttacct	cctttgccat	gcttgattag	caccggtggg	1740
ttgctgaaac	ccaccatttg	tagcgccaca	tcttctcttt	cttcctcgct	gtccacgatt	1800
acctctggtg	atggcgggcg	ctcgggcttg	ggagaagggc	gcttcttttt	cttcttgggc	1860
gcaatggcca	aatccgccgc	cgaggctgat	ggccgcgggc	tgggtgtgcg	cggcaccagc	1920
gcgtcttggt	atgagtcttc	ctcgtcctcg	gactcgatac	gccgcctcat	ccgctttttt	1980
gggggcgccc	ggggaggcgg	cggcgacggg	gacggggacg	acacgtcctc	catggttggg	2040
ggacgtcgcg	ccgcaccgcg	tccgcgctcg	ggggtggttt	cgcgctgctc	ctcttcccga	2100
ctggccatth	ccttctccta	taggcagaaa	aagatcatgt	cgacaacgtt	ggagtcagtc	2160
gagaagaagg	acagcctaac	cgcctcctct	gagttcgcca	ccaccgcctc	caccgatgcc	2220
gccaacgcgc	ctaccacctt	ccccgtcgag	gcacccccgc	ttgaggagga	ggaagtgatt	2280
atcgagcagg	accaggtttt	tgtaagcgaa	gacgacgagg	accgctcagt	accaacagag	2340
gataaaaaagc	aagaccagga	caacgcagag	gcaaacgagg	aacaagtcgg	gcggggggac	2400
gaaaggcatg	gcgactacct	agatgtggga	gacgacgtgc	tgttgaagca	tctgcagcgc	2460
cagtgcgcca	ttatctgcga	cgcgttgcaa	gagcgcagcg	atgtgcccct	cgccatagcg	2520
gatgtcagcc	ttgcctacga	acgccacctt	ttctcaccgc	gcgtaccccc	caaacgccaa	2580
gaaaacggca	catgcgagcc	caacccgcgc	ctcaacttct	accccgattt	tgccgtgcca	2640
gaggtgcttg	ccacctatca	catctttttc	caaaactgca	agatacccct	atcctgccgt	2700
gccaacccga	gccgagcgga	caagcagctg	gccttgccggc	agggcgctgt	catacctgat	2760
atcgccctcg	tcaacgaagt	gccaaaaatc	tttgagggtc	ttggacgcga	cgagaagcgc	2820
gcggcaaacg	ctctgcaaca	ggaaaacagc	gaaaatgaaa	gtcactctgg	agtgttgggtg	2880
gaactcgagg	gtgacaacgc	gcgcctagcc	gtactaaaac	gcagcatcga	ggtcacccac	2940
tttgcctacc	cggcacttaa	cctaccccc	aaggctcatga	gcacagtcac	gagtgagctg	3000
atcgtgcgcc	gtgcgcagcc	cctggagagg	gatgcaaatt	tgcaagaaca	aacagaggag	3060
ggcctacccg	cagttggcga	cgagcagcta	gcgcgctggc	ttcaaacgcg	cgagcctgcc	3120
gacttggagg	agcgacgcaa	actaatgatg	gccgcagtgc	tcgttaccgt	ggagcttgag	3180
tgcatgcagc	ggttctttgc	tgaccgggag	atgcagcgca	agctagagga	aacattgcac	3240
tacacctttc	gacagggtta	cgtacgccag	gcctgcaaga	tctccaacgt	ggagctctgc	3300
aacctggtct	cctaccttgg	aattttgcac	gaaaaccgcc	ttgggcaaaa	cgtgcttcat	3360
tccacgctca	agggcgaggc	gcgcgcgac	tacgtccgcg	actgcgttta	cttatttcta	3420
tgctacacct	ggcagacggc	catgggcgtt	tggcagcagt	gcttggagga	gtgcaacctc	3480

aaggagctgc	agaaactgct	aaagcaaaac	ttgaaggacc	tatggacggc	cttcaacgag	3540
cgctccgtgg	ccgcgcacct	ggcggacatc	atcccccccg	aacgcctgct	taaaaccctg	3600
caacagggtc	tgccagactt	caccagtcaa	agcatgttgc	agaactttag	gaactttatc	3660
ctagagcgct	caggaatctt	gcccgcaccc	tgtgtgacac	ttcctagcga	ctttgtgccc	3720
attaagtacc	gcgaatgccc	tccgcgcgtt	tggggccact	gctaccttct	gcagctagcc	3780
aactaccttg	cctaccactc	tgacataatg	gaagacgtga	gcggtgacgg	tctactggag	3840
tgtcactgtc	gctgcaacct	atgcaccccg	caccgctccc	tggtttgcaa	ttcgcagctg	3900
cttaacgaaa	gtcaaattat	cggtagcctt	gagctgcagg	gtccctcgcc	tgacgaaaag	3960
tccgcggctc	cggggttgaa	actcactccg	gggctgtgga	cgtcggctta	ccttcgcaaa	4020
tttgtacctg	aggactacca	cgcacacgag	attaggttct	acgaagacca	atcccgcccg	4080
cctaattgcg	agcttaccgc	ctgcgtcatt	accaggggcc	acattcttgg	ccaattgcaa	4140
gccatcaaca	aagcccgcca	agagtttctg	ctacgaaagg	gacggggggg	ttacttggac	4200
ccccagtccg	gcgaggagct	caacccaatc	ccccgcgcgc	cgcagcccta	tcagcagcag	4260
ccgcggggcc	ttgcttccca	ggatggcacc	caaaaagaag	ctgcagctgc	cgccgccacc	4320
cacggacgag	gaggaatact	gggacagtca	ggcagaggag	gttttgacg	aggaggagga	4380
ggacatgatg	gaagactggg	agagcctaga	cgaggaagct	tccgaggtcg	aagaggtgtc	4440
agacgaaaca	ccgtcaccct	cggctgcatt	cccctcgccg	gcgcccaga	aatcggaac	4500
cggttccagc	atggctacaa	cctccgctcc	tcaggcgccg	ccggcactgc	ccgttcgccc	4560
acccaaccgt	agatgggaca	ccactggaac	cagggccggg	aagtccaagc	agccgccgcc	4620
gttagcccaa	gagcaacaac	agcgccaagg	ctaccgctca	tggcgcgggc	acaagaacgc	4680
catagttgct	tgcttgcaag	actgtggggg	caacatctcc	ttcgcccgcc	gctttcttct	4740
ctaccatcac	ggcgtggcct	tccccgtaa	catcctgcat	tactaccgtc	atctctacag	4800
cccatactgc	accggcggca	gcggcagcaa	cagcagcggc	cacacagaag	caaaggcgac	4860
cggatagcaa	gactctgaca	aagcccaaga	aatccacagc	ggcggcagca	gcaggaggag	4920
gagcgctgcg	tctggcgccc	aacgaaccgc	tatcgaccgc	cgagcttaga	aacaggattt	4980
ttcccactct	gtatgctata	tttcaacaga	gcagggggcca	agaacaagag	ctgaaaataa	5040
aaaacaggtc	tctgcgatcc	ctcaccgcga	gctgcctgta	tcacaaagtc	gacagcgaag	5100
atcagcttcg	gcgcacgctg	gaagacgcgg	aggctctctt	cagtaaatac	tgcgcgctga	5160
ctcttaagga	ctagtttcgc	gccctttctc	aaatttaagc	gcgaaaacta	cgtcatctcc	5220
agcgttcgaa	cttacacttt	ttcatacatt	gcccaagaat	aaagaatcgt	ttgtgttatg	5280
tttcaacgtg	tttatttttc	aattgcagaa	aatttcaagt	catttttcat	tcagtagtat	5340

agccccacca	ccacatagct	tatacagatc	accgtacctt	aatcaaactc	acagaaccct	5400
agtattcaac	ctgccacctc	cctcccaaca	cacagagtac	acagtccttt	ctccccggct	5460
ggccttaaaa	agcatcatat	catgggtaac	agacatattc	ttaggtgtta	tattccacac	5520
ggtttcctgt	cgagccaaac	gctcatcagt	gatattaata	aactccccgg	gcagctcact	5580
taagttcatg	tcgctgtcca	gctgctgagc	cacaggctgc	tgtccaactt	gcggttgctt	5640
aacgggcggc	gaaggagaag	tccacgccta	catgggggta	gagtcataat	cgtgcatcag	5700
gatagggcgg	tggtgctgca	gcagcgcgcg	aataaactgc	tgccgccgcc	gctccgtcct	5760
gcaggaatac	aacatggcag	tggtctcctc	agcgatgatt	cgcaccgccc	gcagcataag	5820
gcgccttgtc	ctccgggcac	agcagcgcac	cctgatctca	cttaaatacag	cacagtaact	5880
gcagcacagc	accacaatat	tgttcaaaaat	cccacagtgc	aaggcgctgt	atccaaagct	5940
catggcgggg	accacagaac	ccacgtggcc	atcataccac	aagcgcgagg	agattaagtg	6000
gcgacccctc	ataaacacgc	tggacataaa	cattacctct	tttggcatgt	tgtaattcac	6060
cacctcccg	taccatataa	acctctgatt	aaacatggcg	ccatccacca	ccatcctaaa	6120
ccagctggcc	aaaacctgcc	cgccggctat	acactgcagg	gaaccgggac	tggaacaatg	6180
acagtggaga	gcccaggact	cgtaaccatg	gatcatcatg	ctcgtcatga	tatcaatggt	6240
ggcacaacac	aggcacacgt	gcatacactt	cctcaggatt	acaagctcct	cccgcgttag	6300
aaccatatcc	cagggaacaa	ccatttcctg	aatcagcgta	aatcccacac	tgcaagggaag	6360
acctcgcacg	taactcacgt	tgtgcattgt	caaagtgtta	cattcgggca	gcagcggatg	6420
atcctccagt	atggtagcgc	gggtttctgt	ctcaaaagga	ggtagacgat	ccctactgta	6480
cggagtgcgc	cgagacaacc	gagatcgtgt	tggtcgtagt	gtcatgccaa	atggaacgcc	6540
ggacgtagtc	atatttcctg	aagcaaaacc	aggtgcgggc	gtggatccga	caaacagatc	6600
tgctctccg	gtctcgccgc	ttagatcgct	ctgtgtagta	gttgtagtat	atccactctc	6660
tcaaagcatc	caggcgcccc	ctggcttcgg	gttctatgta	aactccttca	tgcgccgctg	6720
ccctgataac	atccaccacc	gcagaataag	ccacaccag	ccaacctaca	cattcgttct	6780
gcgagtcaca	cacgggagga	gcgggaagag	ctggaagaac	catgtttttt	ttttttattc	6840
caaaagatta	tccaaaacct	caaaatgaag	atctattaag	tgaacgcgct	cccctccggt	6900
ggcgtggtca	aactctacag	ccaaagaaca	gataatggca	tttgtaagat	gttgacacaat	6960
ggcttccaaa	aggcaaacgg	ccctcacgtc	caagtggacg	taaaggctaa	acccttcagg	7020
gtgaatctcc	tctataaaca	ttccagcacc	ttcaaccatg	cccaaataat	tctcatctcg	7080
ccaccttctc	aatatatctc	taagcaaatc	ccgaatatta	agtccggcca	ttgtaaaaat	7140
ctgctccaga	gcgccctcca	ccttcagcct	caagcagcga	atcatgattg	caaaaattca	7200
ggttcctcac	agacctgtat	aagattcaaa	agcggaacat	taacaaaaat	accgcgatcc	7260

cgtaggtccc	ttcgcagggc	cagctgaaca	taatcgtgca	ggtctgcacg	gaccagcgcg	7320
gccacttccc	cgccaggaac	catgacaaaa	gaacccacac	tgattatgac	acgcatactc	7380
ggagctatgc	taaccagcgt	agccccgatg	taagcttggt	gcatgggcgg	cgatataaaa	7440
tgcaaggtgc	tgctcaaaaa	atcaggcaaa	gcctcgcgca	aaaaagaaag	cacatcgtag	7500
tcatgctcat	gcagataaag	gcaggttaagc	tccggaacca	ccacagaaaa	agacaccatt	7560
tttctctcaa	acatgtctgc	gggtttctgc	ataaacacaa	aataaaataa	caaaaaaaca	7620
tttaaacatt	agaagcctgt	cttacaacag	gaaaaacaac	ccttataagc	ataagacgga	7680
ctacggccat	gctggcgtga	ccgtaaaaaa	actggtcacc	gtgattaataa	agcaccaccg	7740
acagctcctc	ggtcatgtcc	ggagtcataa	tgtaagactc	ggtaaacaca	tcaggttgat	7800
tcacatcggt	cagtgtctaa	aagcgaccga	aatagcccg	gggaatacat	acccgcaggc	7860
gtagagacaa	cattacagcc	cccataggag	gtataacaaa	attaatagga	gagaaaaaca	7920
cataaacacc	tgaaaaaccc	tcctgcctag	gcaaaatagc	accctcccgc	tccagaacaa	7980
catacagcgc	ttccacagcg	gcagccataa	cagtcagcct	taccagtaaa	aaagaaaacc	8040
tattaaaaaa	acaccactcg	acacggggat	cccaccagct	caatcagtca	cagtgtaaaa	8100
aagggccaag	tgacagagca	gtatatatag	gactaaaaaa	tgacgtaacg	gttaaagtcc	8160
acaaaaaaca	cccagaaaac	cgcacgcgaa	cctacgcccc	gaaacgaaag	ccaaaaaacc	8220
cacaacttcc	tcaaatacgtc	acttccggtt	tcccacgtta	cgtcacttcc	cattttaaga	8280
ttaattaaac	tacaattccc	aacacataca	agttactccg	ccctaaaacc	tacgtcaccc	8340
gc						8342

<210> 5
 <211> 35934
 <212> ADN
 <213> Adenovirus de tipo 5
 <400> 5

ES 2 928 689 T3

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt	60
ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtgtg gcggaagtgt	120
gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg	180
gtgtgcgccg gtgtacacag gaagtgacaa ttttcgcgcg gttttaggcg gatgttgtag	240
taaatttggg cgtaaccgag taagatttgg ccattttcgc gggaaaactg aataagagga	300
agtgaaatct gaataatttt gtgttactca tagcgcgtaa tatttgtcta gggccgcggg	360
gactttgacc gtttacgtgg agactcgccc aggtgttttt ctcaggtgtt ttccgcgttc	420
cgggtcaaag ttggcgtttt attattatag tcagctgacg tgtagtgtat ttatacccg	480
tgagttcctc aagaggccac tcttgagtgc cagcgagtag agttttctcc tccgagccgc	540

tccgacaccg	ggactgaaaa	tgagacatat	tatctgccac	ggaggtgtta	ttaccgaaga	600
aatggccgcc	agtcttttgg	accagctgat	cgaagaggtta	ctggctgata	atcttccacc	660
tcctagccat	tttgaaccac	ctacccttca	cgaactgtat	gatttagacg	tgacggcccc	720
cgaagatccc	aacgaggagg	cggtttcgca	gatttttccc	gactctgtaa	tgttggcggg	780
gcaggaaggg	attgacttac	tcacttttcc	gccggcgccc	ggttctccgg	agccgcctca	840
cctttcccgg	cagcccagagc	agccggagca	gagagccttg	ggtccggttt	ctatgccaaa	900
ccttgatccg	gaggtgatcg	atcttacctg	ccacgaggct	ggctttccac	ccagtgcgca	960
cgaggatgaa	gaggggtgagg	agtttgtgtt	agattatgtg	gagcaccctg	ggcacgggtg	1020
caggtcttgt	cattatcacc	ggaggaatac	gggggaccca	gatattatgt	gttcgctttg	1080
ctatatgagg	acctgtggca	tgtttgtcta	cagtaagtga	aaattatggg	cagtgggtga	1140
tagagtgggtg	ggtttgggtg	ggtaattttt	tttttaattt	ttacagtttt	gtggtttaaa	1200
gaattttgta	ttgtgatttt	tttaaaaggt	cctgtgtctg	aacctgagcc	tgagcccagag	1260
ccagaaccgg	agcctgcaag	acctaccgcg	cgtcctaaaa	tggcgccctg	tatcctgaga	1320
cgcccagacat	cacctgtgtc	tagagaatgc	aatagtagta	cggatagctg	tgactccggg	1380
ccttctaaca	cacctcctga	gatacaccgc	gtgggtccgc	tgtgccccat	taaaccagtt	1440
gccgtgagag	ttgggtgggcg	tcgccaggct	gtggaatgta	tcgaggactt	gcttaacgag	1500
cctgggcaac	ctttggactt	gagctgtaaa	cgccccaggc	cataaggtgt	aaacctgtga	1560
ttgcgtgtgt	ggttaacgcc	tttgtttgct	gaatgagttg	atgtaagttt	aataaagggg	1620
gagataatgt	ttaacttgca	tggcgtgtta	aatggggcgg	ggcttaaagg	gtatataatg	1680
cgccgtgggc	taatcttggg	tacatctgac	ctcatggagg	cttgggagtg	tttggagat	1740
ttttctgctg	tgcgtaactt	gctggaacag	agctctaaca	gtacctcttg	gttttgagg	1800
tttctgtggg	gctcatccca	ggcaaagtta	gtctgcagaa	ttaaggagga	ttacaagtgg	1860
gaatttgaag	agcttttgaa	atcctgtggg	gagctgtttg	attctttgaa	tctgggtcac	1920
caggcgcttt	tccaagagaa	ggtcatcaag	actttggatt	tttccacacc	ggggcgcgct	1980
gcggctgctg	ttgctttttt	gagttttata	aaggataaat	ggagcgaaga	aaccatctg	2040
agcggggggg	acctgctgga	ttttctggcc	atgcatctgt	ggagagcggg	tgtgagacac	2100
aagaatcgcc	tgctactgtt	gtcttcogtc	cgcccggcga	taataccgac	ggaggagcag	2160
cagcagcagc	aggaggaagc	caggcgggcg	cggcaggagc	agagcccatg	gaacccgaga	2220
gccggcctgg	accctcgggg	atgaatgttg	tacaggtggc	tgaactgtat	ccagaactga	2280
gacgcatttt	gacaattaca	gaggatgggc	aggggctaaa	gggggtaaag	agggagcggg	2340
gggcttggtg	ggctacagag	gaggctagga	atctagcttt	tagcttaatg	accagacacc	2400

gtcctgagtg	tattactttt	caacagatca	aggataattg	cgctaagag	cttgatctgc	2460
tggcgcagaa	gtattccata	gagcagctga	ccacttactg	gctgcagcca	ggggatgatt	2520
ttgaggaggc	tattagggtg	tatgcaaagg	tggcacttag	gccagattgc	aagtacaaga	2580
tcagcaaaact	tgtaaatatc	aggaattggt	gctacatttc	tgggaacggg	gccgaggtgg	2640
agatagatac	ggaggatagg	gtggccttta	gatgtagcat	gataaatatg	tggccggggg	2700
tgcttggcat	ggacgggggtg	gttattatga	atgtaagggt	tactggcccc	aatttttagcg	2760
gtacggtttt	cctggccaat	accaacctta	tctacacggg	tgtaagcttc	tatgggttta	2820
acaatacctg	tgtggaagcc	tggaccgatg	taagggttcg	gggctgtgcc	ttttactgct	2880
gctggaaggg	ggtggtgtgt	cgccccaaaa	gcagggcttc	aattaagaaa	tgcctctttg	2940
aaaggtgtac	cttgggtatc	ctgtctgagg	gtaactccag	ggtgcgccac	aatgtggcct	3000
ccgactgtgg	ttgcttcatg	ctagtgaaaa	gcgtggctgt	gattaagcat	aacatggtat	3060
gtggcaactg	cgaggacagg	gcctctcaga	tgtgacctg	ctcggacggc	aactgtcacc	3120
tgtgaagac	cattcacgta	gccagccact	ctcgcaaggc	ctggccagtg	tttgagcata	3180
acatactgac	ccgctgttcc	ttgcatttgg	gtaacaggag	gggggtgttc	ctaccttacc	3240
aatgcaatth	gagtcacact	aagatattgc	ttgagcccga	gagcatgtcc	aaggtgaacc	3300
tgaacgggggt	gtttgacatg	accatgaaga	tctggaagggt	gctgaggtag	gatgagaccc	3360
gcaccaggtg	cagaccctgc	gagtgtggcg	gtaaacatat	taggaaccag	cctgtgatgc	3420
tggatgtgac	cgaggagctg	aggcccgatc	acttgggtgct	ggcctgcacc	cgcgctgagt	3480
ttggctctag	cgatgaagat	acagattgag	gtactgaaat	gtgtgggcgt	ggcttaaggg	3540
tgggaaagaa	tatataaggt	gggggtctta	tgtagtthttg	tatctgtthtt	gcagcagccg	3600
ccgcccgcct	gagcaccaac	togthttgatg	gaagcattgt	gagctcatat	ttgacaacgc	3660
gcatgcccc	atgggcccggg	gtgcgtcaga	atgtgatggg	ctccagcatt	gatggtcgcc	3720
ccgtcctgcc	cgcaaactct	actaccttga	cctacgagac	cgtgtctgga	acgccgttgg	3780
agactgcagc	ctccgcccgc	gcttcagccg	ctgcagccac	cgcccgcggg	attgtgactg	3840
actthtgctth	cctgagccccg	cttgcaagca	gtgcagcttc	ccgttcatcc	gcccgcgatg	3900
acaagttgac	ggctctthttg	gcacaattgg	attctthtgac	ccgggaactt	aatgtcgthtt	3960
ctcagcagct	gttggatctg	cgccagcagg	thttctgccct	gaaggctthcc	thccctccca	4020
atgcggtthta	aaacataaat	aaaaaaccag	actctgthttg	gattthggatc	aagcaagtgt	4080
cttgctgtct	thattthaggg	gtthttgcgcg	cgcggtaggc	ccgggaccag	cggtctcggt	4140
cgttgaggggt	cctgtgtatt	thttccagga	cgtggtaaaag	gtgactctgg	atgttcagat	4200
acatgggcat	aagcccgtct	ctgggggtgga	ggtagcacca	ctgcagagct	tcattgctgcg	4260
gggtggtgtt	gtagatgatc	cagtcgtagc	aggagcgctg	ggcgtggtgc	ctaaaaatgt	4320

ctttcagtag	caagctgatt	gccaggggca	ggcccttggg	gtaagtgttt	acaaagcggg	4380
taagctggga	tgggtgcata	cgtggggata	tgagatgcat	cttggactgt	atTTTTtaggt	4440
tggctatggt	cccagccata	tccctccggg	gattcatggt	gtgcagaacc	accagcacag	4500
tgtatccggg	gcacttggga	aatttgtcat	gtagcttaga	aggaaatgcg	tggaagaact	4560
tggagacgcc	cttgtgacct	ccaagatTTT	ccatgcattc	gtccataatg	atggcaatgg	4620
gcccacgggc	ggcggcctgg	gcgaagatat	ttctgggata	actaacgtca	tagttgtggt	4680
ccaggatgag	atcgtcatag	gccatTTTTa	caaagcgagg	gcgagggggtg	ccagactgcg	4740
gtataatggg	tccatccggc	ccagggggcg	agttaccctc	acagatttgc	atttcccacg	4800
ctttgagttc	agatgggggg	atcatgtcta	cctgcggggc	gatgaagaaa	acggtttccg	4860
gggtagggga	gatcagctgg	gaagaaagca	ggttcctgag	cagctgcgac	ttaccgcagc	4920
cgggtggccc	gtaaatcaca	cctattaccg	gctgcaactg	gtagttaaga	gagctgcagc	4980
tgccgtcatc	cctgagcagg	ggggccactt	cgttaagcat	gtccctgact	cgcattgttt	5040
ccctgaccaa	atccgccaga	aggcgctcgc	cgcccagcga	tagcagttct	tgcaaggaag	5100
caaagTTTT	caacggtttg	agaccgtccg	ccgtaggcat	gcttttgagc	gtttgaccaa	5160
gcagttccag	gcgggtccac	agctcgggtc	cctgctctac	ggcatctcga	tccagcatat	5220
ctcctcggtt	cgcggggttg	ggcggttttc	gctgtacggc	agtagtcggg	gctcgtccag	5280
acggggcagg	gtcatgtctt	tccacggggc	caggggtcct	gtcagcgtag	tctgggtcac	5340
gggtgaaggg	tgcgctccgg	gctgcgcgct	ggccaggggtg	cgcttgaggc	tggtcctgct	5400
gggtgctgaag	cgtgcgcggg	cttcgccctg	cgcgtcgggc	aggtagcatt	tgaccatggg	5460
gtcatagtcc	agcccctccg	cggcgtggcc	cttggcgcg	agcttgccct	tggaggaggc	5520
gccgcacgag	gggcagtgca	gacttttgag	ggcgtagagc	ttgggcgcga	gaaataccga	5580
ttccggggag	taggcattcc	cgcgcaggc	cccgcagacg	gtctcgcat	ccacgagcca	5640
gggtgagctct	ggccgttcgg	ggtcaaaaac	caggtttccc	ccatgctttt	tgatgcgttt	5700
cttacctctg	gtttccatga	gccgggtgtc	acgctcgggtg	acgaaaaggc	tgtccgtgtc	5760
cccgtataca	gacttgagag	gcctgtcctc	gagcgggtgt	ccgcggtcct	cctcgtatag	5820
aaactcggac	cactctgaga	caaaggctcg	cgtccaggcc	agcacgaagg	aggctaagtg	5880
ggaggggtag	cggtcgttgt	ccactagggg	gtccactcgc	tccaggggtgt	gaagacacat	5940
gtcgccctct	tcggcatcaa	ggaagggtgat	tggttttag	gtgtaggcca	cgtgaccggg	6000
tgttcctgaa	ggggggctat	aaaagggggg	gggggcgcgt	tcgtcctcac	tctcttccgc	6060
atcgctgtct	gcgagggcca	gctgttgggg	tgagtactcc	ctctgaaaag	cgggcatgac	6120
ttctgcgcta	agattgtcag	tttccaaaaa	cgaggaggat	ttgatattca	cctggcccgc	6180

ggtgatgcct ttgaggggtgg ccgcatccat ctggtcagaa aagacaatct ttttgttgct	6240
aagcttggtg gcaaacgacc cgtagagggc gttggacagc aacttggcga tggagcgag	6300
ggtttggttt ttgtcgcgat cggcgcgctc cttggccgcg atgtttagct gcacgtattc	6360
gcgcgcaacg caccgccatt cgggaaagac ggtggtgcgc tcgtcgggca ccaggtgcac	6420
gcgccaaccg cggttggtgca gggtgacaag gtcaacgctg gtggctacct ctccgcgtag	6480
gcgctcgttg gtccagcaga ggcggccgcc cttgcgcgag cagaatggcg gtagggggctc	6540
tagctgcgtc tcgtccgggg ggtctgcgtc cagggtaaag accccgggca gcaggcgcgc	6600
gtcgaagtag tctatcttgc atccttgcaa gtctagcgcc tgctgccatg cgcgggcggc	6660
aagcgcgcgc tcgtatgggt tgagtggggg accccatggc atgggggtggg tgagcgcgga	6720
ggcgtacatg ccgcaaattg cgtaaacgta gaggggctct ctgagtattc caagatatgt	6780
agggtagcat cttccaccgc ggatgctggc gcgcacgtaa tcgtatagtt cgtgcgaggg	6840
agcgaggagg tcgggaccga ggttgctacg ggcgggctgc tctgctcgga agactatctg	6900
cctgaagatg gcatgtgagt tggatgatat ggttggacgc tgggaagacgt tgaagctggc	6960
gtctgtgaga cctaccgcgt caccgacgaa ggaggcgtag gagtgcgcga gcttgttgac	7020
cagctcggcg gtgacctgca cgtctagggc gcagtagtcc agggtttcct tgatgatgtc	7080
atacttatcc tgtccctttt ttttccacag ctcgcggttg aggacaaact cttcgcggtc	7140
tttccagtag tcttggtatg gaaaccgcgc ggcctccgaa cggtaagagc ctagcatgta	7200
gaactggttg acggcctggt aggcgcagca tcccttttct acgggtagcg cgtatgcctg	7260
cgcggccttc cggagcgagg tgtgggtgag cgcaaagggtg tccctgacca tgactttgag	7320
gtactggtat ttgaagtcag tgtcgtcgca tccgccctgc tcccagagca aaaagtccgt	7380
gcgctttttg gaacgcggat ttggcagggc gaagggtgaca tcgttgaaga gtatctttcc	7440
cgcgcgagggc ataaagttgc gtgtgatgcg gaagggtccc ggcacctcg aacggttggt	7500
aattacctgg gcggcgagca cgatctcgtc aaagccggtg atgttggtggc ccacaatgta	7560
aagttccaag aagcgcggga tgcccttgat ggaaggcaat tttttaagtt cctcgtaggt	7620
gagctcttca ggggagctga gcccggtgctc tgaaagggcc cagtctgcaa gatgagggtt	7680
ggaagcgacg aatgagctcc acaggtcacg ggccattagc atttgcaggt ggtcgcgaaa	7740
ggtcctaacc tggcgacctg tggccatttt ttctgggggtg atgcagtaga aggtaaagcg	7800
gtcttggtcc cagcgggtccc atccaagggt cgcgggctagg tctcgcgcgg cagtcactag	7860
aggctcatct ccgccgaact tcatgaccag catgaagggc acgagctgct tcccaaaggc	7920
ccccatccaa gtataggtct ctacatcgta ggtgacaaag agacgctcg tgcgaggatg	7980
cgagccgatc gggaagaact ggatctcccg ccaccaattg gaggagtggc tattgatgtg	8040
gtgaaagtag aagtccctgc gacggggcga acactcgtgc tggcttttgt aaaaacgtgc	8100

gcagtactgg	cagcgggtgca	cgggctgtac	atcctgcacg	aggttgacct	gacgaccgcg	8160
cacaaggaag	cagagtggga	atttgagccc	ctcgctggc	gggtttggct	ggtggtcttc	8220
tacttcggct	gcttgtcctt	gaccgtctgg	ctgctcgagg	ggagttacgg	tggatcggac	8280
caccacgccg	cgcgagccca	aagtccagat	gtccgcgcgc	ggcggtcgga	gcttgatgac	8340
aacatcgcgc	agatgggagc	tgtccatggt	ctggagctcc	cgcggcgtca	ggtcaggcgg	8400
gagctcctgc	aggtttacct	cgcatagacg	ggtcagggcg	cgggctagat	ccaggtgata	8460
cctaattttcc	aggggctggt	tgggtggcggc	gtcgatggct	tgcaagaggc	cgcattccccg	8520
cggcgcgact	acggtaccgc	gcggcggggcg	gtgggcgcgc	ggggtgtcct	tggatgatgc	8580
atctaaaagc	ggtgacgcgg	gcgagcccc	ggaggtaggg	ggggctccgg	acccgccggg	8640
agagggggca	ggggcacgtc	ggcgccgcgc	gcgggcagga	gctggtgctg	cgcgcgtagg	8700
ttgctggcga	acgcgacgac	gcggcgggtt	atctcctgaa	tctggcgcct	ctgcgtgaag	8760
acgacggggc	cggtgagctt	gaacctgaaa	gagagttcga	cagaatcaat	ttcgggtgtcg	8820
ttgacggcgg	cctggcgcaa	aatctcctgc	acgtctcctg	agttgtcttg	ataggcgatc	8880
tcggccatga	actgctcgat	ctcttcctcc	tggagatctc	cgcgtccggc	tcgctccacg	8940
gtggcggcga	ggtcgttgga	aatgcggggc	atgagctgcg	agaaggcgtt	gaggcctccc	9000
tcgttccaga	cgcggctgta	gaccacgccc	ccttcggcat	cgcggggcgcg	catgaccacc	9060
tgcgcgagat	tgagctccac	gtgccggggc	aagacggcgt	agtttcgcag	gcgctgaaag	9120
aggtagttga	gggtggtggc	ggtgtgttct	gccacgaaga	agtacataac	ccagcgtcgc	9180
aacgtggatt	cgttgatata	ccccaaaggc	tcaaggcgct	ccatggcctc	gtagaagtcc	9240
acggcgaagt	tgaaaaactg	ggagttgcgc	gccgacacgg	ttaactcctc	ctccagaaga	9300
cggatgagct	cggcgacagt	gtcgcgcacc	tcgcgctcaa	aggctacagg	ggcctcttct	9360
tcttcttcaa	tctcctcttc	cataagggcc	tccccttctt	cttcttcttg	cggcgggtggg	9420
ggagggggga	cacggcggcg	acgacggcgc	accgggaggc	ggtcgacaaa	gcgctcgatc	9480
atctccccgc	ggcgacggcg	catggtctcg	gtgacggcgc	ggcgttctc	gcgggggcgc	9540
agttggaaga	cgcgcgccgt	catgtcccgg	ttatgggttg	gcggggggct	gccatgcggc	9600
agggatacgg	cgctaacgat	gcattctcaac	aattgttgtg	taggtactcc	gccgccgagg	9660
gacctgagcg	agtcgcgcatc	gaccggatcg	gaaaacctct	cgagaaaggc	gtctaaccag	9720
tcacagtcgc	aaggtaggct	gagcacctgt	gcgggcggca	gcgggcggcg	gtcgggggtg	9780
tttctggcgg	aggtgctgct	gatgatgtaa	ttaaagtagg	cggctcttgag	acggcggatg	9840
gtcgacagaa	gcaccatgtc	cttgggtccg	gcctgctgaa	tcgcgaggcg	gtcggccatg	9900
ccccaggctt	cgttttgaca	tcggcgcagg	tctttgtagt	agtcttgcat	gagcctttct	9960

accggcactt	cttcttctcc	ttcctcttgt	cctgcatctc	ttgcatctat	cgctgcggcg	10020
gcggcgaggt	ttggccgtag	gtggcgccct	cttctctcca	tgcgtgtgac	cccgaagccc	10080
ctcatcggct	gaagcagggc	taggtcggcg	acaacgcgct	cggctaatat	ggcctgctgc	10140
acctgcgtga	gggtagactg	gaagtcatcc	atgtccacaa	agcgggtgga	tgcgcccgtg	10200
ttgatggtgt	aagtgcagtt	ggccataacg	gaccagttaa	cggctctggtg	acccggctgc	10260
gagagctcgg	tgtacctgag	acgcgagtaa	gccctcgagt	caaatacgta	gtcgttgcaa	10320
gtccgcacca	ggtactggta	tcccacaaaa	aagtgcggcg	gcggctggcg	gtagaggggc	10380
cagcgtaggg	tggccggggc	tccggggggc	agatcttcca	acataaggcg	atgatatccg	10440
tagatgtacc	tggacatcca	ggtgatgccg	gcggcggtgg	tggaggcgcg	cggaaagtcg	10500
cggacgcggt	tccagatgtt	gcgcagcggc	aaaaagtgct	ccatggtcgg	gacgctctgg	10560
ccggtcaggc	gcgcgcaatc	gttgacgctc	tagaccgtgc	aaaaggagag	cctgtaagcg	10620
ggcactcttc	cgtggtctgg	tggataaatt	cgcaagggtta	tcatggcgga	cgaccgggggt	10680
tcgagccccg	tatccggccg	tccgccgtga	tccatgcggt	taccgcccgc	gtgtcgaacc	10740
caggtgtgcg	acgtcagaca	acgggggaggt	gtcctctttg	gcttccttcc	aggcgcgggc	10800
gctgctgcgc	tagctttttt	ggccactggc	cgcgcgcagc	gtaagcggtt	aggctggaaa	10860
gcgaaagcat	taagtggctc	gctccctgta	gccggagggt	tattttccaa	gggttgagtc	10920
gcgggacccc	cggttcgagt	ctcggaaccg	ccggactgcg	gcgaacgggg	gtttgcctcc	10980
ccgtcatgca	agaccccgtc	tgcaaattcc	tccggaaaca	gggacgagcc	ccttttttgc	11040
ttttcccaga	tgcattccgt	gctgcggcag	atgcgcccc	ctcctcagca	gcggcaagag	11100
caagagcagc	ggcagacatg	cagggcaccc	tcccctcctc	ctaccgcgtc	aggaggggcg	11160
acatccgcgg	ttgacgcggc	agcagatggt	gattacgaac	ccccgcggcg	ccgggcccgg	11220
cactacctgg	acttgaggga	gggcgagggc	ctggcgcggc	taggagcgcc	ctctcctgag	11280
cggcacccaa	gggtgcagct	gaagcgtgat	acgcgtgagg	cgtacgtgcc	gcggcagaac	11340
ctgtttcgcg	accgcgaggg	agaggagccc	gaggagatgc	gggatcgaaa	gttccacgca	11400
gggcgcgagc	tgcggcatgg	cctgaatcgc	gagcggttgc	tgcgcgagga	ggactttgag	11460
cccgacgcgc	gaaccgggat	tagtcccgcg	cgcgcacacg	tggcggccgc	cgacctggta	11520
accgcatacg	agcagacggt	gaaccaggag	attaactttc	aaaaaagctt	taacaaccac	11580
gtgcgtacgc	ttgtggcgcg	cgaggagggtg	gctataggac	tgatgcatct	gtgggacttt	11640
gtaagcgcg	tggagcaaaa	cccaaatagc	aagccgctca	tggcgagct	gttccttata	11700
gtgcagcaca	gcagggacaa	cgaggcattc	agggatgcgc	tgctaaacat	agtagagccc	11760
gagggcccgt	ggctgctcga	tttgataaac	atcctgcaga	gcatagtggt	gcaggagcgc	11820
agcttgagcc	tggctgacaa	ggtggccgcc	atcaactatt	ccatgcttag	cctgggcaag	11880

ttttacgccc	gcaagatata	ccatacccct	tacgttccca	tagacaagga	ggtaaagatc	11940
gaggggttct	acatgcgcac	ggcgctgaag	gtgcttacct	tgagcgacga	cctgggcggt	12000
tatcgcaacg	agcgcatcca	caaggccgtg	agcgtgagcc	ggcggcgcgga	gctcagcgac	12060
cgcgagctga	tgcacagcct	gcaaagggcc	ctggctggca	cgggcagcgg	cgatagagag	12120
gccgagtcct	actttgacgc	gggcgctgac	ctgcgctggg	ccccaagccg	acgcgccttg	12180
gaggcagctg	gggcccggacc	tgggctggcg	gtggcacccg	cgcgcgctgg	caacgtcggc	12240
ggcgtggagg	aatatgacga	ggacgatgag	tacgagccag	aggacggcga	gtactaagcg	12300
gtgatgtttc	tgatcagatg	atgcaagacg	caacggaccc	ggcgggtgcg	gcggcgctgc	12360
agagccagcc	gtccggcctt	aactccacgg	acgactggcg	ccaggtcacg	gaccgcatca	12420
tgtcgctgac	tgcgcgcaat	cctgacgcgt	tccggcagca	gccgcaggcc	aaccggctct	12480
ccgcaattct	ggaagcgggt	gtcccggcgc	gcgcaaacc	cacgcacgag	aaggtgctgg	12540
cgatcgtaaa	cgcgctggcc	gaaaacaggg	ccatccggcc	cgacgaggcc	ggcctggtct	12600
acgacgcgct	gcttcagcgc	gtggctcggt	acaacagcgg	caacgtgcag	accaacctgg	12660
accggctggt	gggggatgtg	cgcgaggccg	tggcgcagcg	tgagcgcgcg	cagcagcagg	12720
gcaacctggg	ctccatgggt	gcactaaacg	ccttcctgag	tacacagccc	gccaacgtgc	12780
cgcgggggaca	ggaggactac	accaactttg	tgagcgcact	gcggtctaat	gtgactgaga	12840
caccgcaaag	tgagggtgtac	cagtctgggc	cagactatct	tttcagacc	agtagacaag	12900
gcctgcagac	cgtaaacctg	agccaggctt	tcaaaaactt	gcaggggctg	tgggggggtgc	12960
gggctcccac	aggcgaccgc	gcgaccgtgt	ctagcttgct	gacgcccac	tcgcgcctgt	13020
tgctgctgct	aatagcgcgc	ttcacggaca	gtggcagcgt	gtcccgggac	acatacctag	13080
gtcacttgct	gacactgtac	cgcgaggcca	taggtcaggc	gcatgtggac	gagcatactt	13140
tccaggagat	tacaagtgtc	agccgcgcgc	tggggcagga	ggacacgggc	agcctggagg	13200
caaccctaaa	ctacctgctg	accaaccggc	ggcagaagat	cccctcggtg	cacagtttaa	13260
acagcgagga	ggagcgcatt	ttgcgctacg	tgcagcagag	cgtgagcctt	aacctgatgc	13320
gcgacggggg	aacgcccagc	gtggcgctgg	acatgaccgc	gcgcaacatg	gaaccgggca	13380
tgtatgcctc	aaaccggccg	tttatcaacc	gcctaattga	ctacttgcat	cgcgcggccg	13440
ccgtgaaccc	cgagtatttc	accaatgcca	tcttgaaccc	gcactggcta	ccgccccttg	13500
gtttctacac	cgggggattc	gaggtgcccg	agggtaacga	tggattcctc	tgggacgaca	13560
tagacgacag	cgtgttttcc	ccgcaaccgc	agaccctgct	agagttgcaa	cagcgcgagc	13620
aggcagaggc	ggcgctgcga	aaggaaagct	tccgcaggcc	aagcagcttg	tccgatctag	13680
gcgctgcggc	cccgcggtca	gatgctagta	gcccatttcc	aagcttgata	gggtctctta	13740

ES 2 928 689 T3

ccagcactcg	caccacccgc	ccgcgcctgc	tgggcgagga	ggagtaccta	aacaactcgc	13800
tgctgcagcc	gcagcgcgaa	aaaaacctgc	ctccggcatt	tccaacaac	gggatagaga	13860
gcctagtgga	caagatgagt	agatggaaga	cgtacgcgca	ggagcacagg	gacgtgccag	13920
gcccgcgccc	gcccacccgt	cgtcaaaggc	acgaccgtca	gcggggtctg	gtgtgggagg	13980
acgatgactc	ggcagacgac	agcagcgtcc	tggatttggg	agggagtggc	aacccgtttg	14040
cgcaccttcg	ccccaggctg	gggagaatgt	tttaaaaaaa	aaaaaagcat	gatgcaaaat	14100
aaaaaaactca	ccaaggccat	ggcaccgagc	gttgggttttc	ttgtattccc	cttagtatgc	14160
ggcgcgcggc	gatgtatgag	gaaggtcctc	ctccctccta	cgagagtgtg	gtgagcgcgg	14220
cgccagtggc	ggcggcgctg	ggttctccct	tcgatgctcc	cctggaccgc	ccgtttgtgc	14280
ctccgcggta	cctgcggcct	accgggggga	gaaacagcat	ccgttactct	gagttggcac	14340
ccctattcga	caccacccgt	gtgtacctgg	tggacaacaa	gtcaacggat	gtggcatccc	14400
tgaactacca	gaacgaccac	agcaactttc	tgaccacggt	cattcaaaac	aatgactaca	14460
gcccggggga	ggcaagcaca	cagaccatca	atcttgacga	ccggtcgcac	tggggcggcg	14520
acctgaaaac	catcctgcat	accaacatgc	caaagtgtgaa	cgagtccatg	tttaccaata	14580
agtttaaggc	gcgggtgatg	gtgtcgcgct	tgcctactaa	ggacaatcag	gtggagctga	14640
aatacgagtg	ggtggagttc	acgctgcccg	agggcaacta	ctccgagacc	atgaccatag	14700
accttatgaa	caacgcgata	gtggagcaact	acttgaaagt	gggcagacag	aacgggggttc	14760
tggaaagcga	catcggggta	aagtttgaca	cccgcaactt	cagactgggg	tttgacccccg	14820
tcaactggtct	tgtcatgcct	ggggtatata	caaacgaagc	cttccatcca	gacatcattt	14880
tgctgccagg	atgcgggggtg	gacttcaccc	acagccgcct	gagcaacttg	ttgggcatcc	14940
gcaagcggca	acccttccag	gagggcttta	ggatcaccta	cgatgatctg	gaggggtgga	15000
acattcccgc	actggtggat	gtggacgcct	accaggcgag	cttgaaagat	gacaccgaac	15060
agggcggggg	tggcgcaggc	ggcagcaaca	gcagtggcag	cggcgcgga	gagaactcca	15120
acgcggcagc	cgcggcaatg	cagccgggtg	aggacatgaa	cgatcatgcc	attcgcggcg	15180
acacctttgc	cacacgggct	gaggagaagc	gcgctgaggc	cgaagcagcg	gccgaagctg	15240
ccgccccccg	tgcgcaaccc	gaggtcgaga	agcctcagaa	gaaaccgggtg	atcaaacc	15300
tgacagagga	cagcaagaaa	cgcagttaca	acctaataag	caatgacagc	accttcaccc	15360
agtaccgcag	ctggtacctt	gcatacaact	acggcgaccc	tcagaccgga	atccgctcat	15420
ggaccctgct	ttgcactcct	gacgtaacct	gcggctcgga	gcaggtctac	tggtcgttgc	15480
cagacatgat	gcaagacccc	gtgaccttcc	gctccacgcg	ccagatcagc	aactttccgg	15540
tgggtgggcgc	cgagctgttg	cccgtgcact	ccaagagctt	ctacaacgac	caggccgtct	15600
actcccaact	catccgccag	tttacctctc	tgaccacagt	gttcaatcgc	tttcccgaga	15660

accagatttt	ggcgcgccc	ccagccccc	ccatcaccac	cgtcagtga	aacgttcctg	15720
ctctcacaga	tcacgggacg	ctaccgctgc	gcaacagcat	cggaggagtc	cagcgagtga	15780
ccattactga	cgccagacgc	cgcacctgcc	cctacgttta	caaggccctg	ggcatagtct	15840
cgcgcgcgt	cctatcgagc	cgcacttttt	gagcaagcat	gtccatcctt	atatcgccca	15900
gcaataacac	aggctggggc	ctgcgcttcc	caagcaagat	gtttggcggg	gccaagaagc	15960
gctccgacca	acacccagtg	cgcgtgcgcg	ggcactaccg	cgcgccctgg	ggcgcgcaca	16020
aacgcggccg	cactgggcgc	accaccgtcg	atgacgccat	cgcgcgggtg	gtggaggagg	16080
cgcgcaacta	cacgcccacg	ccgccaccag	tgtccacagt	ggacgcggcc	attcagaccg	16140
tgggtgcgcg	agcccggcgc	tatgctaaaa	tgaagagacg	gcggaggcgc	gtagcacgtc	16200
gccaccgccc	ccgaccgggc	actgccgccc	aacgcgcggc	ggcggccctg	cttaaccgcg	16260
cacgtcgcac	cggccgacgg	gcggccatgc	gggcgcgtcg	aaggctggcc	gcgggtattg	16320
tcactgtgcc	ccccagggtc	aggcgacgag	cggccgcgcg	agcagccgcg	gccattagtg	16380
ctatgactca	gggtcgcagg	ggcaacgtgt	attgggtgcg	cgactcggtt	agcggcctgc	16440
gcgtgcccgt	gcgcacccgc	ccccgcgcga	actagattgc	aagaaaaaac	tacttagact	16500
cgtactgttg	tatgtatcca	gcggcggcgg	cgcgcaacga	agctatgtcc	aagcgcaaaa	16560
tcaaagaaga	gatgctccag	gtcatcgcgc	cggagatcta	tggccccccg	aagaaggaag	16620
agcaggatta	caagccccga	aagctaaagc	gggtcaaaaa	gaaaaagaaa	gatgatgatg	16680
atgaacttga	cgacgaggtg	gaactgctgc	acgctaccgc	gcccaggcga	cgggtacagt	16740
ggaaaggtcg	acgcgtaaaa	cgtgttttgc	gaccgggcac	caccgtagtc	tttacgcccg	16800
gtgagcgctc	caccgcgacc	tacaagcgcg	tgtatgatga	ggtgtacggc	gacgaggacc	16860
tgcttgagca	ggccaacgag	cgcctcgggg	agtttgccta	cggaaagcgg	cataaggaca	16920
tgctggcggt	gccgctggac	gagggcaacc	caacacctag	cctaaagccc	gtaacactgc	16980
agcaggtgct	gcccgcgctt	gcaccgtccg	aagaaaagcg	cggcctaaag	cgcgagtctg	17040
gtgacttggc	accacccgtg	cagctgatgg	tacccaagcg	ccagcgactg	gaagatgtct	17100
tggaaaaaat	gaccgtggaa	cctgggctgg	agcccagaggt	ccgcgtgcgg	ccaatcaagc	17160
aggtggcgcc	gggactgggc	gtgcagaccg	tggacgttca	gatacccact	accagtagca	17220
ccagtattgc	caccgccaca	gagggcatgg	agacacaaac	gtccccgggt	gcctcagcgg	17280
tggcggatgc	cgcgggtgcag	gcggtcgctg	cggccgcgctc	caagacctct	acggaggtgc	17340
aaacggaccc	gtggatgttt	cgcgtttcag	ccccccggcg	cccgcgccgt	tcgaggaagt	17400
acggcgccgc	cagcgcgcta	ctgccogaat	atgccctaca	tccttccatt	gcgcctaccc	17460
ccggctatcg	tggctacacc	taccgcccc	gaagacgagc	aactaccgga	cgccgaacca	17520

ccactggaac	ccgccgccgc	cgtcgccgtc	gccagcccgt	gctggccccg	atttccgtgc	17580
gcaggggtggc	tcgcgaagga	ggcaggaccc	tggtgctgcc	aacagcgcg	taccacccca	17640
gcatcgttta	aaagccggtc	tttgtggttc	ttgcagatat	ggccctcacc	tgccgcctcc	17700
gtttccccgt	gccgggattc	cgaggaagaa	tgcaccgtag	gaggggcatg	gccggccacg	17760
gcctgacggg	cggcatgcgt	cgtgcgcacc	accggcgggc	gcgcgcgtcg	caccgtcgca	17820
tgcgcggcgg	tatcctgccc	ctccttattc	cactgatcgc	cgcggcgatt	ggcgccgtgc	17880
ccggaattgc	atccgtggcc	ttgcaggcgc	agagacactg	attaaaaaca	agttgcatgt	17940
ggaaaaatca	aaataaaaag	tctggactct	cacgctcgct	tggtcctgta	actattttgt	18000
agaatggaag	acatcaactt	tgcgtctctg	gccccgcgac	acggctcgcg	cccgttcattg	18060
ggaaactggc	aagatatcgg	caccagcaat	atgagcgggtg	gcgccttcag	ctggggctcg	18120
ctgtggagcg	gcattaaaaa	tttcggttcc	accgttaaga	actatggcag	caaggcctgg	18180
aacagcagca	caggccagat	gctgagggat	aagttgaaag	agcaaaattt	ccaacaaaag	18240
gtggtagatg	gcctggcctc	tggcattagc	ggggtggtgg	acctggccaa	ccaggcagtg	18300
caaaataaga	ttaacagtaa	gcttgatccc	cgcctcccgc	tagaggagcc	tccaccggcc	18360
gtggagacag	tgtctccaga	ggggcgtggc	gaaaagcgtc	cgcgccccga	cagggaagaa	18420
actctggtga	cgcaaataga	cgagcctccc	tcgtacgagg	aggcactaaa	gcaaggcctg	18480
cccaccaccc	gtcccacgc	gcccattggc	accggagtg	tgggccagca	cacaccctga	18540
acgctggacc	tgcctcccc	cgcgcacacc	cagcagaaac	ctgtgctgcc	aggcccgacc	18600
gccgttggtg	taaccctg	tagccgcgcg	tcctgcgcc	gcgcgcgcag	cggctccgcga	18660
tcgttgcggc	ccgtagccag	tggcaactgg	caaagcacac	tgaacagcat	cgtgggtctg	18720
ggggtgcaat	ccctgaagcg	ccgacgatgc	ttctgatagc	taacgtgtcg	tatgtgtgtc	18780
atgtatgcgt	ccatgtcgcc	gccagaggag	ctgctgagcc	gccgcgcgcc	cgttttccaa	18840
gatggctacc	ccttcgatga	tgcgcagtg	gtcttacatg	cacatctcgg	gccaggacgc	18900
ctcgagtagc	ctgagccccg	ggctggtgca	gtttgccccg	gccaccgaga	cgtacttcag	18960
cctgaataac	aagtttagaa	acccacgggt	ggcgccctacg	cacgacgtga	ccacagaccg	19020
gtcccagcgt	ttgacgctgc	ggttcatccc	tgtggaccgt	gaggatactg	cgtactcgta	19080
caaggcgcg	ttcacccctag	ctgtgggtga	taaccgtgtg	ctggacatgg	cttcacgta	19140
ctttgacatc	cgcggcgctgc	tggacagggg	ccctactttt	aagccctact	ctggcactgc	19200
ctacaacgcc	ctggctccca	agggtgcccc	aaatccttgc	gaatgggatg	aagctgctac	19260
tgctcttgaa	ataaacctag	aagaagagga	cgatgacaac	gaagacgaag	tagacgagca	19320
agctgagcag	caaaaaactc	acgtatttgg	gcaggcgcc	tattctggta	taaatattac	19380
aaaggagggt	attcaaatag	gtgtcgaagg	tcaaacacct	aaatatgccg	ataaaacatt	19440

tcaacctgaa cctcaaatag gagaatctca gtggtacgaa acagaaatta atcatgcagc	19500
tgggagagtc ctaaaaaaga ctacccaat gaaaccatgt tacggttcat atgcaaaacc	19560
cacaaatgaa aatggagggc aaggcattct tgtaaagcaa caaatggaa agctagaaag	19620
tcaagtggaa atgcaatddd tctcaactac tgaggcagcc gcaggcaatg gtgataactt	19680
gactcctaaa gtggtattgt acagtgaaga tgtagatata gaaaccccag aactcatat	19740
ttcttacatg cccactatta aggaaggtaa ctcacgagaa ctaatgggcc aacaatctat	19800
gccaacagg cctaattaca ttgcttttag ggacaatddd attggtctaa tgtattacaa	19860
cagcacgggt aatatgggtg ttctggcggg ccaagcatcg cagttgaatg ctgtttaga	19920
tttgcaagac agaaacacag agctttcata ccagcttttg cttgattcca ttggtgatag	19980
aaccaggtag ttttctatgt ggaatcaggc tgttgacagc tatgatccag atgttagaat	20040
tattgaaaat catggaactg aagatgaact tccaaattac tgctttccac tgggaggtgt	20100
gattaataca gagactctta ccaaggtaaa acctaaaaca ggtcaggaaa atggatggga	20160
aaaagatgct acagaatddd cagataaaaa tgaaataaga gttggaaata attttgccat	20220
ggaaatcaat ctaaatgcca acctgtggag aaatttcctg tactccaaca tagcgctgta	20280
tttgcccagc aagctaaagt acagtccttc caacgtaaaa atttctgata acccaaacac	20340
ctacgactac atgaacaagc gagtggtagc tcccggtcta gtggactgct acattaacct	20400
tggagcacgc tgggtcccttg actatatgga caacgtcaac ccatttaacc accaccgcaa	20460
tgctggcctg cgctaccgct caatgttgct gggcaatggt cgctatgtgc ccttccacat	20520
ccagggtgct cagaagttct ttgccattaa aaacctcctt ctctgcccgg gctcatacac	20580
ctacgagtgg aacttcagga aggatgttaa catggttctg cagagctccc taggaaatga	20640
cctaaggggt gacggagcca gcattaagtt tgatagcatt tgcctttacg ccaccttctt	20700
ccccatggcc cacaacaccg cctccacgct tgaggccatg cttagaaacg acaccaacga	20760
ccagtccttt aacgactatc tctccgccc caacatgctc taccctatac ccgccaacgc	20820
taccaacgtg cccatatcca tcccctccc caactgggag gctttccgag gctgggcctt	20880
cacgcgcctt aagactaagg aaacccatc actgggctcg ggctacgacc cttattacac	20940
ctactctggc tctataacct acctagatgg aaccttttac ctcaaccaca ctttaagaa	21000
ggtggccatt acctttgact cttctgtcag ctggcctggc aatgaccgcc tgcttacctc	21060
caacgagttt gaaattaagc gctcagttga cggggaggggt tacaacgttg cccagtgtaa	21120
catgaccaa gactggttcc tgggtacaaat gctagctaac tataacattg gctaccaggg	21180
cttctatatc ccagagagct acaaggacc catgtactcc ttctttagaa acttccagcc	21240
catgagccgt cagggtggtg atgatactaa atacaaggac taccaacagg tgggcatcct	21300

acaccaaacac	aacaactctg	gatttggttg	ctaccttgcc	cccaccatgc	gcgaaggaca	21360
ggcctaccct	gctaacttcc	cctatccgct	tataggcaag	accgcagttg	acagcattac	21420
ccagaaaaag	tttctttgcg	atcgcacccct	ttggcgcatc	ccattctcca	gtaactttat	21480
gtccatgggc	gcactcacag	acctggggcca	aaaccttctc	tacgccaaact	ccgcccacgc	21540
gctagacatg	acttttgagg	tggatcccat	ggacgagccc	acccttcttt	atgttttggt	21600
tgaagtcttt	gacgtggtcc	gtgtgcacca	gccgcaccgc	ggcgtcatcg	aaaccgtgta	21660
cctgcgcacg	cccttctcgg	ccggcaacgc	cacaacataa	agaagcaagc	aacatcaaca	21720
acagctgccg	ccatgggctc	cagtgcagcag	gaactgaaag	ccattgtcaa	agatcttggg	21780
tgtggggccat	atTTTTTggg	cacctatgac	aagcgctttc	caggctttgt	ttctccacac	21840
aagctcgcct	gcgccatagt	caatacggcc	ggtcgcgaga	ctggggggcg	acactggatg	21900
gcctttgcct	ggaacccgca	ctcaaaaaca	tgtacctct	ttgagccctt	tggttttct	21960
gaccagcgac	tcaagcaggt	ttaccagttt	gagtacgagt	cactcctgcg	ccgtagcgcc	22020
attgcttctt	cccccgaccg	ctgtataacg	ctggaaaagt	ccacccaaag	cgtacagggg	22080
cccaactcgg	ccgcctgtgg	actattctgc	tgcattgttc	tccacgcctt	tgccaactgg	22140
ccccaaactc	ccatggatca	caaccccacc	atgaacctta	ttaccggggg	acccaactcc	22200
atgctcaaca	gtccccaggt	acagcccacc	ctgcgtcgca	accaggaaca	gctctacagc	22260
ttcctggagc	gccactcgcc	ctacttccgc	agccacagtg	cgcagattag	gagcgccact	22320
tctttttgtc	acttgaaaaa	catgtaaaaa	taatgtacta	gagacacttt	caataaaggc	22380
aaatgctttt	atttgtagac	tctcgggtga	ttatttacc	ccacccttgc	cgtctgcgcc	22440
gtttaaaaat	caaaggggtt	ctgcgcgcga	tcgctatgcg	ccactggcag	ggacacgttg	22500
cgatactggg	gttttagtgct	ccacttaaac	tcaggcacaa	ccatccgcgg	cagctcgggtg	22560
aagttttcac	tccacaggct	gcgcaccatc	accaacgcgt	ttagcaggtc	gggcgccgat	22620
atcttgaagt	cgcagttggg	gcctccgccc	tgcgcgcgcg	agttgcgata	cacaggggtg	22680
cagcactgga	acactatcag	cgccgggtgg	tgcacgctgg	ccagcacgct	cttgtcggag	22740
atcagatccg	cgtccaggct	ctccgcgttg	ctcagggcga	acggagtcaa	ctttggtagc	22800
tgccttccca	aaaagggcgc	gtgccccaggc	tttgagttgc	actcgcaccg	tagtggcatc	22860
aaaagggtgac	cgtgcccggg	ctgggcgtta	ggatacagcg	cctgcataaa	agccttgatc	22920
tgccttaaaag	ccacctgagc	ctttgcgcct	tcagagaaga	acatgccgca	agacttgccg	22980
gaaaactgat	tggccgggaca	ggccgcgctg	tgcacgcagc	accttgcgtc	ggtgttgag	23040
atctgcacca	catttcgggc	ccaccgggtc	ttcacgatct	tggccttgct	agactgctcc	23100
ttcagcgcgc	gctgcccgtt	tctcgtcgtc	acatccattt	caatcacgtg	ctccttattt	23160
atcataatgc	ttccgtgtag	acacttaagc	tgccttcga	tctcagcgca	gcggtgcagc	23220

cacaacgcgc agcccgtggg ctcgtgatgc ttgtaggtca cctctgcaaa cgactgcagg	23280
tacgcctgca ggaatcgccc catcatcgtc acaaaggtct tgttgctggt gaaggtcagc	23340
tgcaacccgc ggtgctcctc gttcagccag gtcttgcata cggccgccag agcttccact	23400
tggtcaggca gtagtttgaa gttcgccttt agatcgttat ccacgtggta cttgtccatc	23460
agcgcgcgcg cagcctccat gcccttctcc cacgcagaca cgatcggcac actcagcggg	23520
ttcatcaccg taattttcact ttccgcttcg ctgggctctt cctcttcctc ttgcgtccgc	23580
ataccacgcg ccaactgggtc gtcttcattc agccgccgca ctgtgcgctt acctccttg	23640
ccatgcttga ttagcaccgg tgggttgctg aaaccaccca tttgtagcgc cacatcttct	23700
ctttcttcct cgctgtccac gattacctct ggtgatggcg ggcgctcggg cttgggagaa	23760
gggcgcttct ttttcttctt gggcgcaatg gccaaatccg ccgccgaggt cgatggccgc	23820
gggctgggtg tgcgcggcac cagcgcgtct tgtgatgagt cttcctcgtc ctcggaactcg	23880
atacgccgcc tcatccgctt ttttgggggc gcccggggag gcgccggcga cggggacggg	23940
gacgacacgt cctccatggt tgggggacgt cgcgcgcgcac cgcgtccgcg ctcgggggtg	24000
gtttcgcgct gtcctcttc ccgactggcc atttccttct cctataggca gaaaaagatc	24060
atggagtcag tcgagaagaa ggacagccta accgccccct ctgagttcgc caccaccgcc	24120
tccaccgatg ccgccaaacgc gcctaccacc tccccgtcg aggcaccccc gcttgaggag	24180
gaggaagtga ttatcgagca ggaccaggt tttgtaagcg aagacgacga ggaccgctca	24240
gtaccaacag aggataaaaa gcaagaccag gacaacgcag aggcaaacga ggaacaagtc	24300
gggcgggggg acgaaaggca tggcgactac ctagatgtgg gagacgacgt gctgttgaag	24360
catctgcagc gccagtgcgc cattatctgc gacgcgttgc aagagcgcag cgatgtgccc	24420
ctcgccatag cggatgtcag ccttgccctac gaacgccacc tattctcacc gcgcgtaccc	24480
cccaaacgcc aagaaaacgg cacatgcgag cccaacccgc gcctcaactt ctaccccgt	24540
tttgccgtgc cagaggtgct tgccacctat cacatctttt tccaaaactg caagataccc	24600
ctatcctgcc gtgccaaacg cagccgagcg gacaagcagc tggccttgcg gcagggcgct	24660
gtcatacctg atatcgcctc gctcaacgaa gtgccaaaaa tctttgaggg tcttgacgc	24720
gacgagaagc gcgcggcaaa cgctctgcaa caggaaaaca gcgaaaatga aagtcactct	24780
ggagtgttgg tggaaactga gggtgacaac gcgcgcctag ccgtactaaa acgcagcatc	24840
gaggtcaccc actttgccta cccggcactt aacctacccc ccaaggcat gagcacagtc	24900
atgagtgagc tgatcgtgcg ccgtgcgcag cccctggaga gggatgcaaa tttgcaagaa	24960
caaacagagg agggcctacc cgcagttggc gacgagcagc tagcgcgctg gcttcaaacg	25020
cgcgagcctg ccgacttgga ggagcgacgc aaactaatga tggccgcagt gctcgttacc	25080

gtggagcttg	agtgcacgca	gcgggttcttt	gctgacccgg	agatgcagcg	caagctagag	25140
gaaacattgc	actacacctt	tcgacagggc	tacgtacgcc	aggcctgcaa	gatctccaac	25200
gtggagctct	gcaacctggt	ctcctacctt	ggaattttgc	acgaaaaccg	ccttgggcaa	25260
aacgtgcttc	attccacgct	caagggcgag	gcgcgccgcg	actacgtccg	cgactgcgtt	25320
tacttatctt	tatgctacac	ctggcagacg	gccatggggc	tttggcagca	gtgcttggag	25380
gagtgcaccc	tcaaggagct	gcagaaactg	ctaaagcaaa	acttgaagga	cctatggacg	25440
gccttcaacg	agcgctccgt	ggccgcgcac	ctggcggaca	tcattttccc	cgaacgcctg	25500
cttaaaaccc	tgcaacaggg	tctgccagac	ttcaccagtc	aaagcatggt	gcagaacttt	25560
aggaacttta	tcctagagcg	ctcaggaatc	ttgcccggca	cctgctgtgc	acttcctagc	25620
gactttgtgc	ccattaagta	ccgcgaatgc	cctccgccgc	tttggggcca	ctgctacctt	25680
ctgcagctag	ccaactacct	tgccctaccac	tctgacataa	tggaagacgt	gagcgggtgac	25740
ggtctactgg	agtgtcactg	tcgctgcaac	ctatgcaccc	cgcaccgctc	cctgggtttgc	25800
aattcgcagc	tgcttaacga	aagtcaaatt	atcggtagct	ttgagctgca	gggtccctcg	25860
cctgacgaaa	agtccgcggc	tccgggggtt	aaactcactc	cggggctgtg	gacgtcggct	25920
taccttcgca	aatttgtacc	tgaggactac	cacgcccacg	agattaggtt	ctacgaagac	25980
caatcccgcc	cgcctaatac	ggagcttacc	gcctgcgtca	ttaccagagg	ccacattctt	26040
ggccaattgc	aagccatcaa	caaagcccgc	caagagtttc	tgctacgaaa	gggacggggg	26100
gtttacttgg	acccccagtc	cggcgaggag	ctcaacccaa	tccccccgcc	gccgcagccc	26160
tatcagcagc	agccgcgggc	ccttgcttcc	caggatggca	cccaaaaaga	agctgcagct	26220
gccgccgcca	cccacggacg	aggaggaata	ctgggacagt	caggcagagg	aggttttggg	26280
cgaggaggag	gaggacatga	tggaagactg	ggagagccta	gacgaggaag	cttccgaggt	26340
cgaagagggt	tcagacgaaa	caccgtcacc	ctcggtcgca	ttccctcgc	cggcgcccca	26400
gaaatcggca	accggttcca	gcatggctac	aacctccgct	cctcaggcgc	cgccggcact	26460
gcccgttcgc	cgacccaacc	gtagatggga	caccactgga	accagggccg	gtaagtccaa	26520
gcagccgccc	ccgttagccc	aagagcaaca	acagcgccaa	ggctaccgct	catggcgccg	26580
gcacaagaac	gccatagttg	cttgcttgca	agactgtggg	ggcaacatct	ccttcgcccc	26640
ccgctttctt	ctctaccatc	acggcggtgg	cttcccccg	aacatcctgc	attactaccg	26700
tcctctctac	agcccatact	gcaccggcgg	cagcggcagc	aacagcagcg	gccacacaga	26760
agcaaaggcg	accggatagc	aagactctga	caaagcccaa	gaaatccaca	gcggcggcag	26820
cagcaggagg	aggagcgctg	cgtctggcgc	ccaacgaacc	cgtatcgacc	cgcgagctta	26880
gaaacaggat	ttttcccact	ctgtatgcta	tatttcaaca	gagcaggggc	caagaacaag	26940
agctgaaaat	aaaaaacagg	tctctgcgat	ccctcaccgc	cagctgcctg	tatcacaaaa	27000

gcgaagatca gcttcggcgc acgctggaag acgcggaggc tctcttcagt aaatactgcg	27060
cgctgactct taaggactag ttctgcgccc ttctctcaaat ttaagcgcgga aaactacgtc	27120
atctccagcg gccacacccg gcgccagcac ctgttgtcag cgccattatg agcaaggaaa	27180
ttcccacgcc ctacatgtgg agttaccagc cacaaatggg acttgcggtc ggagctgccc	27240
aagactactc aaccogaata aactacatga gcgcgggacc ccacatgata tcccgggtca	27300
acggaatacg cgcccaccga aaccgaattc tcctggaaca ggcggctatt accaccacac	27360
ctcgtaataa ccttaatccc cgtagttggc ccgctgccct ggtgtaccag gaaagtcccg	27420
ctcccaccac tgttggtactt ccagagacg ccaggccga agttcagatg actaactcag	27480
gggcgcagct tgcgggcggc ttctgtcaca ggggtgcggtc gcccgggcag ggtataactc	27540
acctgacaat cagagggcga ggtattcagc tcaacgacga gtcggtgagc tcctcgcttg	27600
gtctccgtcc ggacgggaca ttccagatcg gcggcgccgg ccgctcttca ttcacgcctc	27660
gtcaggcaat cctaactctg cagacctcgt cctctgagcc gcgctctgga ggcattggaa	27720
ctctgcaatt tattgaggag ttgtgtccat cggctctactt taacccttc tcgggacctc	27780
ccggccacta tccgatcaa ttatttcccta actttgacgc ggtaaaggac tcggcggacg	27840
gctacgactg aatgttaagt ggagaggcag agcaactgcg cctgaaacac ctggtccact	27900
gtcgccgcca caagtgttt gccgcggaact ccggtgagtt ttgctacttt gaattgccg	27960
aggatcatat cgagggcccc gcgcacggcg tccggcttac cgccaggga gagcttgccc	28020
gtagcctgat tcgggagttt acccagcgcc ccctgctagt tgagcgggac aggggacctt	28080
gtgttctcac tgtgatttgc aactgtccta accctggatt acatcaagat ctttgttgcc	28140
atctctgtgc tgagtataat aaatacagaa attaaaatat actggggctc ctatcgccat	28200
cctgtaaacy ccaccgtctt caccgcgcca agcaaacc aa ggcgaacctt acctggtact	28260
tttaacatct ctccctctgt gatttacaac agtttcaacc cagacggagt gagtctacga	28320
gagaacctct ccgagctcag ctactccatc agaaaaaaca ccacctcct tacctgccg	28380
gaacgtacga gtgcgtcacc ggccgctgca ccacacctac cgcctgaccg taaaccagac	28440
tttttccgga cagacctcaa taactctgtt taccagaaca ggaggtgagc ttagaaaacc	28500
cttaggggat taggccaaag gcgcagctac tgtgggggtt atgaacaatt caagcaactc	28560
tacgggctat tctaattcag gtttctctag aatcggggtt ggggttattc tctgtcttgt	28620
gattctcttt attcttatac taacgcttct ctgcctaagg ctgcgcgct gctgtgtgca	28680
catttgcat tattgtcagc tttttaaacg ctggggctgc caccacaagat gattaggtac	28740
ataatcctag gtttactcac ccttgogtca gccacggta ccacccaaa ggtggatttt	28800
aaggagccag cctgtaatgt tacattcgca gctgaagcta atgagtgcac cactcttata	28860

aaatgcacca	cagaacatga	aaagctgctt	attcgccaca	aaaacaaaat	tggcaagtat	28920
gctgtttatg	ctatttggca	gccaggtgac	actacagagt	ataatgttac	agttttccag	28980
ggtaaaagtc	ataaaacttt	tatgtatact	tttccatttt	atgaaatgtg	cgacattacc	29040
atgtacatga	gcaaacagta	taagttgtgg	ccccacaaa	attgtgtgga	aaacactggc	29100
actttctgct	gcactgctat	gctaattaca	gtgctcgctt	tggctctgtac	cctactctat	29160
attaaataca	aaagcagacg	cagctttatt	gaggaaaaga	aaatgcctta	atttactaag	29220
ttacaaagct	aatgtcacca	ctaactgctt	tactcgctgc	ttgcaaaaca	aattcaaaaa	29280
gttagcatta	taattagaat	aggattttaa	ccccccggtc	atttcctgct	caataccatt	29340
cccctgaaca	attgactcta	tgtgggatat	gctccagcgc	tacaaccttg	aagtcaggct	29400
tcctggatgt	cagcatctga	ctttggccag	cacctgtccc	gcggatttgt	tccagtccaa	29460
ctacagcgac	ccaccctaac	agagatgacc	aacacaacca	acgcggccgc	cgctaccgga	29520
cttacatcta	ccacaaatac	accccaagtt	tctgcctttg	tcaataactg	ggataacttg	29580
ggcatgtggt	ggttctccat	agcgcttatg	tttgtatgcc	ttattattat	gtggctcatc	29640
tgctgcctaa	agcgcaaacg	cgcccgacca	cccatctata	gtcccatcat	tgtgctacac	29700
ccaaacaatg	atggaatcca	tagattggac	ggactgaaac	acatgttctt	ttctcttaca	29760
gtatgattaa	atgagacatg	attcctcgag	tttttatatt	actgaccctt	gttgcgcttt	29820
ttttgtgcgt	gctccacatt	ggctgcgggt	tctcacatcg	aagtagactg	cattccagcc	29880
ttcacagtct	atttgcttta	cggatttgtc	accctcacgc	tcatctgcag	cctcatcact	29940
gtggtcatcg	cctttatcca	gtgcattgac	tgggtctgtg	tgcgctttgc	atatctcaga	30000
caccatcccc	agtacaggga	caggactata	gctgagcttc	ttagaattct	ttaattatga	30060
aatttactgt	gacttttctg	ctgattatth	gcaccctatc	tgcgttttgt	tccccgacct	30120
ccaagcctca	aagacatata	tcatgcagat	tcaactcgat	atggaatatt	ccaagttgct	30180
acaatgaaaa	aagcgatctt	tccgaagcct	ggttatatgc	aatcatctct	gttatggtgt	30240
tctgcagtac	catcttagcc	ctagctatat	atccctacct	tgacattggc	tggaacgcaa	30300
tagatgccat	gaaccaccca	actttccccg	cgcccgtat	gcttccactg	caacaagttg	30360
ttgccggcgg	ctttgtccca	gccaatcagc	ctcgcccacc	ttctcccacc	cccactgaaa	30420
tcagctactt	taatctaaca	ggaggagatg	actgacaccc	tagatctaga	aatggacgga	30480
attattacag	agcagcgctt	gctagaaaga	cgcagggcag	cggccgagca	acagcgcatg	30540
aatcaagagc	tccaagacat	ggttaacttg	caccagtgc	aaaggggtat	cttttgtctg	30600
gtaaagcagg	ccaaagtcac	ctacgacagt	aataccaccg	gacaccgcct	tagctacaag	30660
ttgccaaacca	agcgtcagaa	attggtggtc	atggtgggag	aaaagcccat	taccataact	30720
cagcactcgg	tagaaaccga	aggctgcatt	cactcacctt	gtcaaggacc	tgaggatctc	30780

tgcaccctta	ttaagaccct	gtgcgggtctc	aaagatctta	ttccctttaa	ctaataaaaa	30840
aaaataataa	agcatcactt	acttaaaatc	agtttagcaaa	tttctgtcca	gtttatttcag	30900
cagcacctcc	ttgccctcct	cccagctctg	gtattgcagc	ttcctcctgg	ctgcaaactt	30960
tctccacaat	ctaaatggaa	tgtcagtttc	ctcctgttcc	tgtccatccg	caccactat	31020
cttcatgttg	ttgcagatga	agcgcgcaag	accgtctgaa	gataccttca	accccggtga	31080
tccatatgac	acggaaaccg	gtcctccaac	tgtgcctttt	cttactcctc	cctttgtatc	31140
ccccaatggg	tttcaagaga	gtccccctgg	ggtactctct	ttgcgcctat	ccgaacctct	31200
agttacctcc	aatggcatgc	ttgcgctcaa	aatgggcaac	ggcctctctc	tggacgaggc	31260
cggcaacctt	acctcccaa	atgtaaccac	tgtgagccca	cctctcaaaa	aaaccaagtc	31320
aaacataaac	ctggaaatat	ctgcaccct	cacagttacc	tcagaagccc	taactgtggc	31380
tgccgcccga	cctctaattg	tcgcgggcaa	cacactcacc	atgcaatcac	aggccccgct	31440
aaccgtgcac	gactccaaac	ttagcattgc	cacccaagga	cccctcacag	tgtcagaagg	31500
aaagctagcc	ctgcaaacat	caggccccct	caccaccacc	gatagcagta	cccttactat	31560
cactgcctca	ccccctctaa	ctactgccac	tggtagcttg	ggcattgact	tgaagagcc	31620
catttataca	caaaatggaa	aactaggact	aaagtacggg	gctcctttgc	atgtaacaga	31680
cgacctaaac	actttgaccg	tagcaactgg	tccaggtgtg	actattaata	atacttcctt	31740
gcaaactaaa	gttactggag	ccttgggttt	tgattcacaa	ggcaatatgc	aacttaatgt	31800
agcaggagga	ctaaggattg	attctcaaaa	cagacgcctt	atacttgatg	ttagttatcc	31860
gtttgatgct	caaaaccaac	taaatctaag	actaggacag	ggccctcttt	ttataaactc	31920
agcccacaac	ttggatatta	actacaacaa	aggcctttac	ttgtttacag	cttcaaacaa	31980
ttccaaaaag	cttgaggtta	acctaaagac	tgccaagggg	ttgatgtttg	acgctacagc	32040
catagccatt	aatgcaggag	atgggcttga	atttggttca	cctaattgcac	caaacacaaa	32100
tcccctcaaa	acaaaaattg	gccatggcct	agaatttgat	tcaaacaagg	ctatggttcc	32160
taaactagga	actggcctta	gttttgacag	cacaggtgcc	attacagtag	gaaacaaaaa	32220
taatgataag	ctaactttgt	ggaccacacc	agctccatct	cctaactgta	gactaaatgc	32280
agagaaagat	gctaaactca	ctttggtctt	aacaaaatgt	ggcagtcaaa	tacttgctac	32340
agtttcagtt	ttggctgtta	aaggcagttt	ggctccaata	tctggaacag	ttcaaagtgc	32400
tcatcttatt	ataagatttg	acgaaaatgg	agtgctacta	aacaattcct	tcctggaccc	32460
agaatattgg	aactttagaa	atggagatct	tactgaaggc	acagcctata	caaacgctgt	32520
tggatttatg	cctaacctat	cagcttatcc	aaaatctcac	ggtaaaactg	ccaaaagtaa	32580
cattgtcagt	caagtttact	taaacggaga	caaaactaaa	cctgtaacac	taaccattac	32640

actaaacggt	acacaggaaa	caggagacac	aactccaagt	gcatactcta	tgtcattttc	32700
atgggactgg	tctggccaca	actacattaa	tgaaatatatt	gccacatcct	cttacacttt	32760
ttcatacatt	gccaagaat	aaagaatcgt	ttgtgttatg	tttcaacgtg	tttatttttc	32820
aattgcagaa	aatttcaagt	catttttcat	tcagtagtat	agccccacca	ccacatagct	32880
tatacagatc	accgtacctt	aatcaaactc	acagaaccct	agtattcaac	ctgccacctc	32940
cctcccaaca	cacagagtac	acagtccttt	ctccccgggt	ggccttaaaa	agcatcatat	33000
catgggtaac	agacatatct	ttaggtgtta	tattccacac	ggtttcctgt	cgagccaaac	33060
gctcatcagt	gatattaata	aactccccgg	gcagctcact	taagttcatg	tcgctgtcca	33120
gctgctgagc	cacaggctgc	tgtccaactt	gcggttgctt	aacgggcggc	gaaggagaag	33180
tccacgccta	catgggggta	gagtcataat	cgtgcatcag	gatagggcgg	tggtgctgca	33240
gcagcgcgcg	aataaactgc	tgccgcgcgc	gctccgtcct	gcaggaatac	aacatggcag	33300
tggtctcctc	agcgatgatt	cgcaccgccc	gcagcataag	gcgccttgtc	ctccgggcac	33360
agcagcgcac	cctgatctca	cttaaatacag	cacagtaact	gcagcacagc	accacaatat	33420
tgttcaaaat	cccacagtgc	aaggcgtgtg	atccaaagct	catggcgggg	accacagaac	33480
ccacgtggcc	atcataccac	aagcgcaggt	agattaagtg	gcgaccctc	ataaacacgc	33540
tggacataaa	cattacctct	tttggcatgt	tgtaattcac	cacctcccgg	taccatataa	33600
acctctgatt	aaacatggcg	ccatccacca	ccatcctaaa	ccagctggcc	aaaacctgcc	33660
cgccggctat	acactgcagg	gaaccgggac	tggaacaatg	acagtggaga	gcccaggact	33720
cgtaaccatg	gatcatcatg	ctcgtcatga	tatcaatgtt	ggcacaacac	aggcacacgt	33780
gcatacactt	cctcaggatt	acaagctcct	cccgcgttag	aaccatatcc	cagggaaaca	33840
cccatctctg	aatcagcgta	aatcccacac	tgcaagggaag	acctcgcacg	taactcacgt	33900
tgtgcattgt	caaagtgtta	cattcgggca	gcagcggatg	atcctccagt	atggtagcgc	33960
gggtttctgt	ctcaaaagga	ggtagacgat	ccctactgta	cggagtgcgc	cgagacaacc	34020
gagatcgtgt	tggtcgtagt	gtcatgccaa	atggaacgcc	ggacgtagtc	atatttcctg	34080
aagcaaaacc	aggtgcgggc	gtgacaaaca	gatctgcgtc	tccggtctcg	ccgcttagat	34140
cgctctgtgt	agtagttgta	gtatatccac	tctctcaaag	catccaggcg	ccccctggct	34200
tcgggttcta	tgtaaaactcc	ttcatgcgcc	gctgccctga	taacatccac	caccgcagaa	34260
taagccacac	ccagccaacc	tacacattcg	ttctgcgagt	cacacacggg	aggagcggga	34320
agagctggaa	gaaccatgtt	tttttttttt	attccaaaag	attatccaaa	acctcaaaat	34380
gaagatctat	taagtgaacg	cgctcccctc	cgggtggcgtg	gtcaaaactct	acagccaaag	34440
aacagataat	ggcatttgta	agatgttgca	caatggcttc	caaaaggcaa	acggccctca	34500
cgtccaagtg	gacgtaaagg	ctaaaccctt	caggggtgaat	ctcctctata	aacattccag	34560

caccttcaac	catgccc aaa	taatttctcat	ctcgccacct	tctcaatata	tctctaagca	34620
aatcccgaat	attaagtccg	gccattgtaa	aaatctgctc	cagagcgccc	tccaccttca	34680
gcctcaagca	gcgaatcatg	attgcaaaaa	ttcaggttcc	tcacagacct	gtataagatt	34740
caaaagcgga	acattaacaa	aaataccgcg	atcccgtagg	tcccttcgca	gggccagctg	34800
aacataatcg	tgcaggtctg	cacggaccag	cgcgggccact	tccccgccag	gaaccatgac	34860
aaaagaaccc	acactgatta	tgacacgcat	actcggagct	atgctaacca	gcgtagcccc	34920
gatgtaagct	tgttgcatgg	gcggcgatat	aaaatgcaag	gtgctgctca	aaaaatcagg	34980
caaagcctcg	cgcaaaaaag	aaagcacatc	gtagtcatgc	tcatgcagat	aaaggcaggt	35040
aagctccgga	accaccacag	aaaaagacac	cattttttctc	tcaaacatgt	ctgcgggttt	35100
ctgcataaac	acaaaataaa	ataacaaaaa	aacattttaa	cattagaagc	ctgtcttaca	35160
acaggaaaaa	caacccttat	aagcataaga	cggactacgg	ccatgccggc	gtgaccgtaa	35220
aaaaactggt	caccgtgatt	aaaaagcacc	accgacagct	cctcggtcac	gtccggagtc	35280
ataatgtaag	actcggtaaa	cacatcaggt	tgattcacat	cggtcagtgc	taaaaagcga	35340
ccgaaatagc	ccgggggaat	acataccgcg	aggcgtagag	acaacattac	agcccccata	35400
ggaggtataa	caaaattaat	aggagagaaa	aacacataaa	cacctgaaaa	accctcctgc	35460
ctaggcaaaa	tagcacctc	ccgctccaga	acaacataca	gcgcttccac	agcggcagcc	35520
ataacagtca	gccttaccag	taaaaaagaa	aacctattaa	aaaaacacca	ctcgacacgg	35580
caccagctca	atcagtcaca	gtgtaaaaaa	gggccaaagt	cagagcgagt	atatatagga	35640
ctaaaaaatg	acgtaacggt	taaagtccac	aaaaaacacc	cagaaaaccg	cacgcgaacc	35700
tacgcccaga	aacgaaagcc	aaaaaaccca	caacttcctc	aaatcgtcac	ttccgttttc	35760
ccacgttacg	tcacttccca	ttttaagaaa	actacaattc	ccaacacata	caagttactc	35820
cgccttaaaa	cctacgtcac	ccgccccgtt	cccacgcccc	gcgccacgtc	acaaactcca	35880
ccccctcatt	atcatattgg	cttcaatcca	aaataaggta	tattattgat	gatg	35934

REIVINDICACIONES

1. Un plásmido auxiliar que comprende al menos un gen rep de virus adenoasociado (AAV) que codifica al menos una proteína Rep de AAV funcional y al menos un gen de virus auxiliar, y que no comprende un gen cap que codifica un conjunto funcional de proteínas Cap.
2. Un sistema de dos plásmidos que comprende el plásmido auxiliar de la reivindicación 1 y un vector plasmídico.
3. El sistema de dos plásmidos de la reivindicación 2, en el que:
 - (i) el vector plasmídico comprende:
 - (a) un gen cap de AAV que codifica al menos una proteína Cap de AAV funcional; o
 - (b) al menos un promotor del gen cap de AAV, un sitio de clonación unido operativamente al promotor del gen cap de AAV y un módulo de expresión flanqueado en al menos un lado por una repetición terminal invertida ("inverted terminal repeat", ITR);
 en el que el vector plasmídico no comprende un gen rep que codifica una proteína Rep funcional y el módulo de expresión comprende un transgén unido operativamente al menos a un elemento de control regulador; y/o
 - (ii) la proporción de plásmido auxiliar a vector plasmídico está entre 1:1 y 1:4.
4. El sistema de dos plásmidos o plásmido auxiliar de cualquiera de las reivindicaciones anteriores, en el que dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 52 funcional, al menos un gen que codifica una proteína Rep 40 funcional y un gen que codifica una proteína Rep 68 funcional.
5. El sistema de dos plásmidos o plásmido auxiliar de cualquiera de las reivindicaciones anteriores, en el que:
 - (i) dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 52 funcional, y el gen que codifica una proteína Rep 52 funcional comprende una secuencia de ácido nucleico que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 1000 nucleótidos de longitud de los nucleótidos 993-2186 de SEQ ID NO:1, o con un tramo de nucleótidos correspondiente en un serotipo diferente de AAV; y/o
 - (ii)
 - (a) dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 40 funcional, y el gen que codifica una proteína Rep 40 funcional comprende una secuencia de ácido nucleico que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 800 nucleótidos de longitud de un tramo de nucleótidos correspondiente a los nucleótidos 993-2252 menos los nucleótidos 1907-2227 de SEQ ID NO:1, o con tramos de nucleótidos correspondientes en un serotipo diferente de AAV; y/o
 - (b) dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 40 funcional, y el gen que codifica una proteína Rep 40 funcional comprende una secuencia de ácido nucleico que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 1100 nucleótidos de longitud de los nucleótidos 993-2252 de SEQ ID NO:1, o con un tramo de nucleótidos correspondiente en un serotipo diferente de AAV; y/o
 - (iii) dicho al menos un gen rep de AAV comprende un gen que codifica una proteína Rep 68 funcional, y el gen que codifica una proteína Rep 68 funcional comprende una secuencia de ácido nucleico que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 1500 nucleótidos de longitud de un tramo de nucleótidos correspondiente a los nucleótidos 321-2252 menos los nucleótidos 1907-2227 de SEQ ID NO:1, o con tramos de nucleótidos correspondientes en un serotipo diferente de AAV.
6. El sistema de dos plásmidos o plásmido auxiliar de cualquiera de las reivindicaciones anteriores, en el que:
 - (i) el plásmido auxiliar no comprende un gen que codifica una proteína Rep 78 funcional; y/o
 - (ii) el plásmido auxiliar no comprende una secuencia contigua de al menos 1800 nucleótidos correspondientes a un tramo de nucleótidos contiguo de longitud equivalente comprendido dentro de los nucleótidos 321-2186 de SEQ ID NO:1 o dentro de un tramo de nucleótidos correspondiente en un serotipo diferente de AAV; y/o
 - (iii) dicho al menos un gen rep de AAV no comprende un promotor p40 interno funcional; y/o
 - (iv) el plásmido auxiliar no comprende un tramo contiguo de la secuencia del gen cap exclusivamente de más de 60 nucleótidos.
7. El sistema de dos plásmidos o plásmido auxiliar de cualquiera de las reivindicaciones anteriores, en el que:
 - (i) el plásmido auxiliar comprende al menos un gen de virus auxiliar y dicho al menos un gen de virus auxiliar comprende un ácido nucleico asociado a virus ("viral associated", VA), un gen E2A y un gen E4; y/o
 - (ii) el plásmido auxiliar comprende al menos un gen de virus auxiliar y dicho al menos un gen de virus auxiliar

comprende un ácido nucleico VA, un gen E2A y un gen E4, y el gen E4 no está ubicado entre el ácido nucleico VA y el gen E2A; y/o

(iii) el plásmido auxiliar comprende al menos un gen de virus auxiliar y el plásmido auxiliar tiene una longitud de entre 12000 pb y 15000 pb; y/o

(iv) el plásmido auxiliar no comprende una secuencia contigua de al menos 3000 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 194-3620 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus; y/o

(v) el plásmido auxiliar no comprende una secuencia contigua de al menos 60 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 4032-4100 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus; y/o

(vi) el plásmido auxiliar no comprende una secuencia contigua de al menos 350 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 4051-4413 de SEQ ID NO:1, o un tramo de nucleótidos correspondiente en un serotipo diferente de AAV.

8. El sistema de dos plásmidos o plásmido auxiliar de cualquiera de las reivindicaciones anteriores, en el que:

(i) el plásmido auxiliar no comprende una secuencia contigua de al menos 600 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 2301-2947 de SEQ ID NO:1, o un tramo de nucleótidos correspondiente en un serotipo diferente de AAV; y/o

(ii) el plásmido auxiliar comprende al menos un gen de virus auxiliar y dicho al menos un gen de virus auxiliar comprende un ácido nucleico VA, un gen E2A y un gen E4, y el ácido nucleico VA, el gen E2A y el gen E4 están comprendidos dentro de una porción contigua de menos de 10000 nucleótidos; y/o

(iii) el plásmido auxiliar no comprende una secuencia contigua de al menos 22000 nucleótidos de un tramo de nucleótidos contiguo de longitud equivalente comprendida dentro de los nucleótidos 10619-32756 de SEQ ID NO:2, o un tramo de nucleótidos correspondiente en un serotipo diferente de adenovirus; y/o

(iv) el plásmido auxiliar y/o el vector plasmídico no comprende un sitio de unión a Rep artificial.

9. El sistema de dos plásmidos o plásmido auxiliar de cualquiera de las reivindicaciones anteriores, en el que el plásmido auxiliar comprende al menos un gen del virus auxiliar y dicho al menos un gen del virus auxiliar está comprendido en un tramo contiguo del plásmido que tiene al menos un 98 % de identidad con la longitud completa o con un fragmento de al menos 8000 nucleótidos de longitud de SEQ ID NO:4.

10. El sistema de dos plásmidos de una cualquiera de las reivindicaciones 2-9, en el que:

(i) el vector plasmídico comprende al menos un promotor del gen cap de AAV, que comprende un promotor p40 de AAV, un promotor p5 y un promotor p19; y/o

(ii) el vector plasmídico no comprende codones de iniciación de la traducción prescindibles; y/o

(iii) el vector plasmídico comprende una región promotora que comprende uno o más promotores, y en la región promotora:

(a) los nucleótidos correspondientes a los nucleótidos 321-323 de SEQ ID NO:1 están ausentes;

(b) los nucleótidos correspondientes a los nucleótidos 766-768 de SEQ ID NO:1 no son ATG;

(c) los nucleótidos correspondientes a los nucleótidos 955-957 de SEQ ID NO:1 están ausentes;

(d) los nucleótidos correspondientes a los nucleótidos 993-995 de SEQ ID NO:1 están ausentes; y

(e) los nucleótidos correspondientes a los nucleótidos 1014-1016 de SEQ ID NO:1 están ausentes; y/o

(iv) el vector plasmídico comprende un esqueleto de menos de 3000 nucleótidos de longitud.

11. El sistema de dos plásmidos de la reivindicación 10(iii)(b), en el que los nucleótidos correspondientes a los nucleótidos 766-768 de SEQ ID NO:1 son ATT.

12. Un método para producir una preparación de AAV recombinante que comprende:

(a) obtener el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico como se define en una cualquiera de las reivindicaciones 1-11;

(b) transfectar una célula huésped con el sistema de dos plásmidos, el plásmido auxiliar o el vector plasmídico como se define en una cualquiera de las reivindicaciones 1-11; y

(c) cultivar la célula huésped en condiciones adecuadas para la producción de AAV recombinante.

13. El método de la reivindicación 12:

(i) que comprende además una etapa de recolección del AAV recombinante para proporcionar una preparación de AAV recombinante; y/o

(ii) en el que la preparación de AAV recombinante comprende un nivel bajo de AAV competente para la replicación (rcAAV); y/o

(iii) en el que el nivel de rcAAV medido es menor que 1 rcAAV en 10^7 AAV recombinantes, menor que 1 rcAAV

en 10^9 AAV recombinantes, o menor que 1 rcAAV en 10^{10} AAV recombinantes; y/o

(iv) en el que la preparación de AAV recombinante comprende un nivel bajo de rcAAV y el nivel de rcAAV es menor que el nivel de rcAAV producido utilizando un método equivalente, excepto que el vector plasmídico comprende al menos un gen rep de AAV y al menos un gen cap de AAV; y/o

5 (v) comprende una etapa de seleccionar una proporción de plásmido auxiliar a vector plasmídico en el que la proporción de plásmido auxiliar a vector plasmídico se selecciona o ajusta a una proporción que permite al usuario obtener la proporción deseada de partículas llenas a totales o el rendimiento alto o deseado del vector AAV recombinante; y/o

10 (vi) en el que la proporción de plásmido auxiliar a vector plasmídico se selecciona o ajusta a una proporción que logra un rendimiento máximo de AAV recombinante con el rendimiento mínimo de partículas vacías alcanzable con dicho rendimiento máximo de AAV recombinante; y/o

(vii) en el que la proporción deseada de partículas llenas a totales es una proporción de partículas llenas a totales que es al menos el 20 % o al menos el 30 % de la proporción de partículas llenas a totales lograda utilizando un método equivalente con una proporción de plásmido auxiliar al vector plasmídico de 1,8:1; y/o

15 (viii) que comprende además una etapa de purificación del AAV recombinante.

Fig. 1

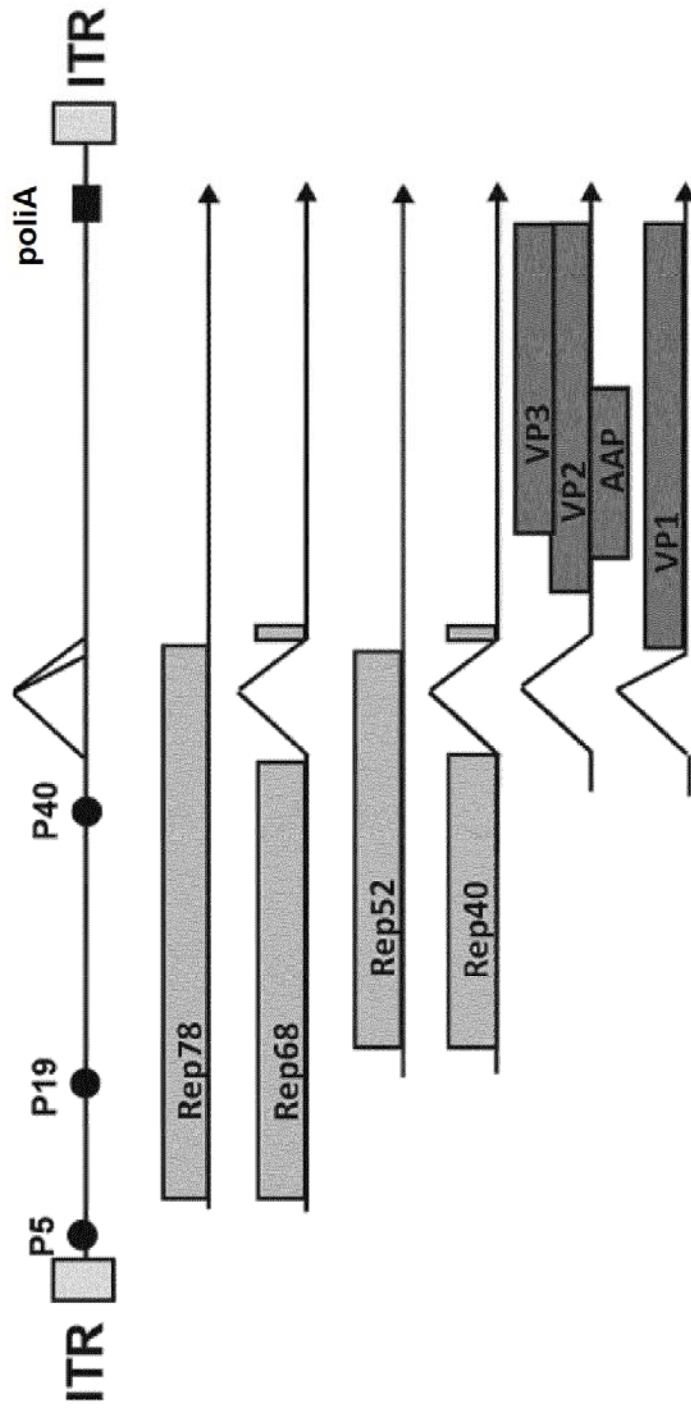


Fig. 2

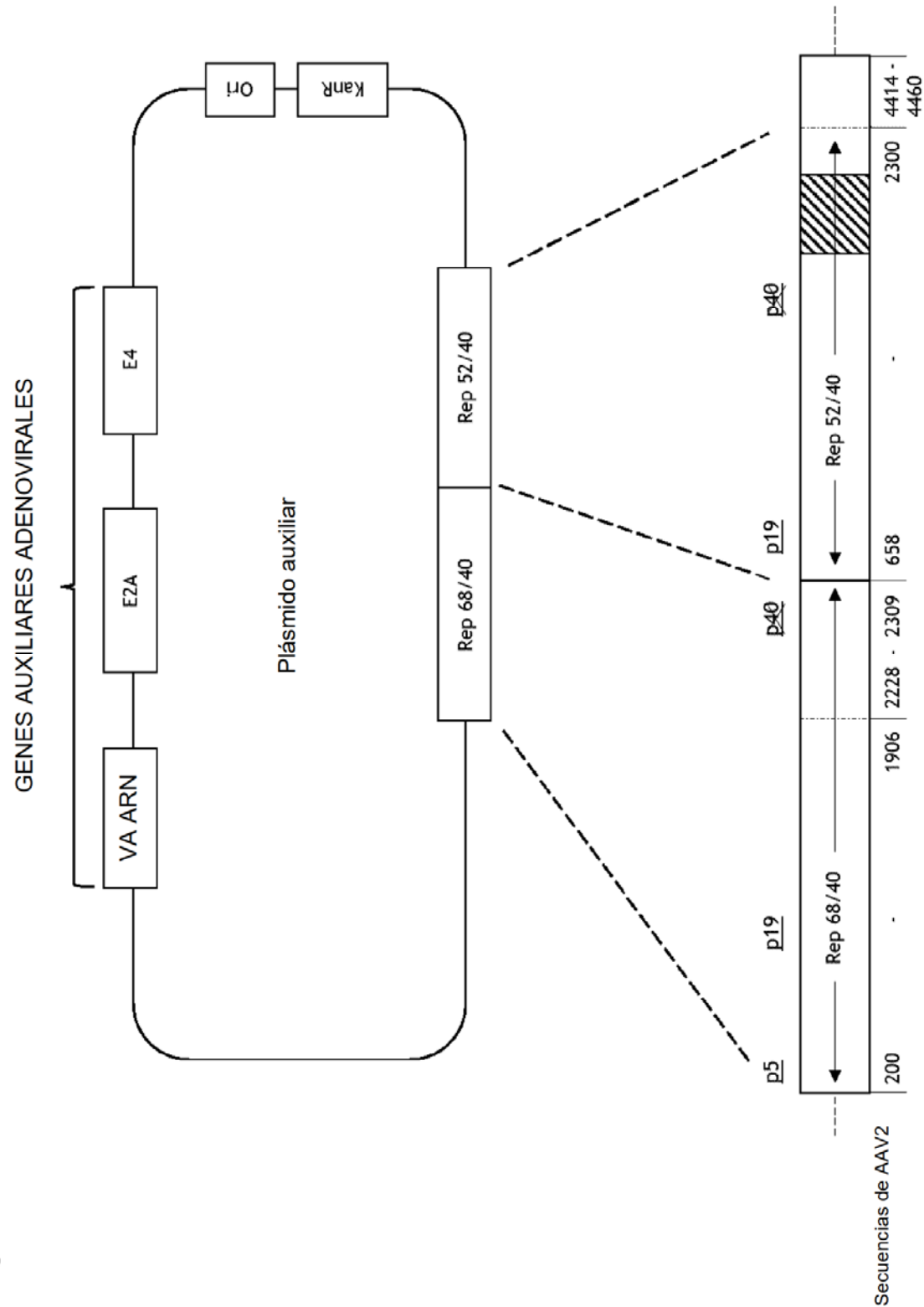


Fig. 3

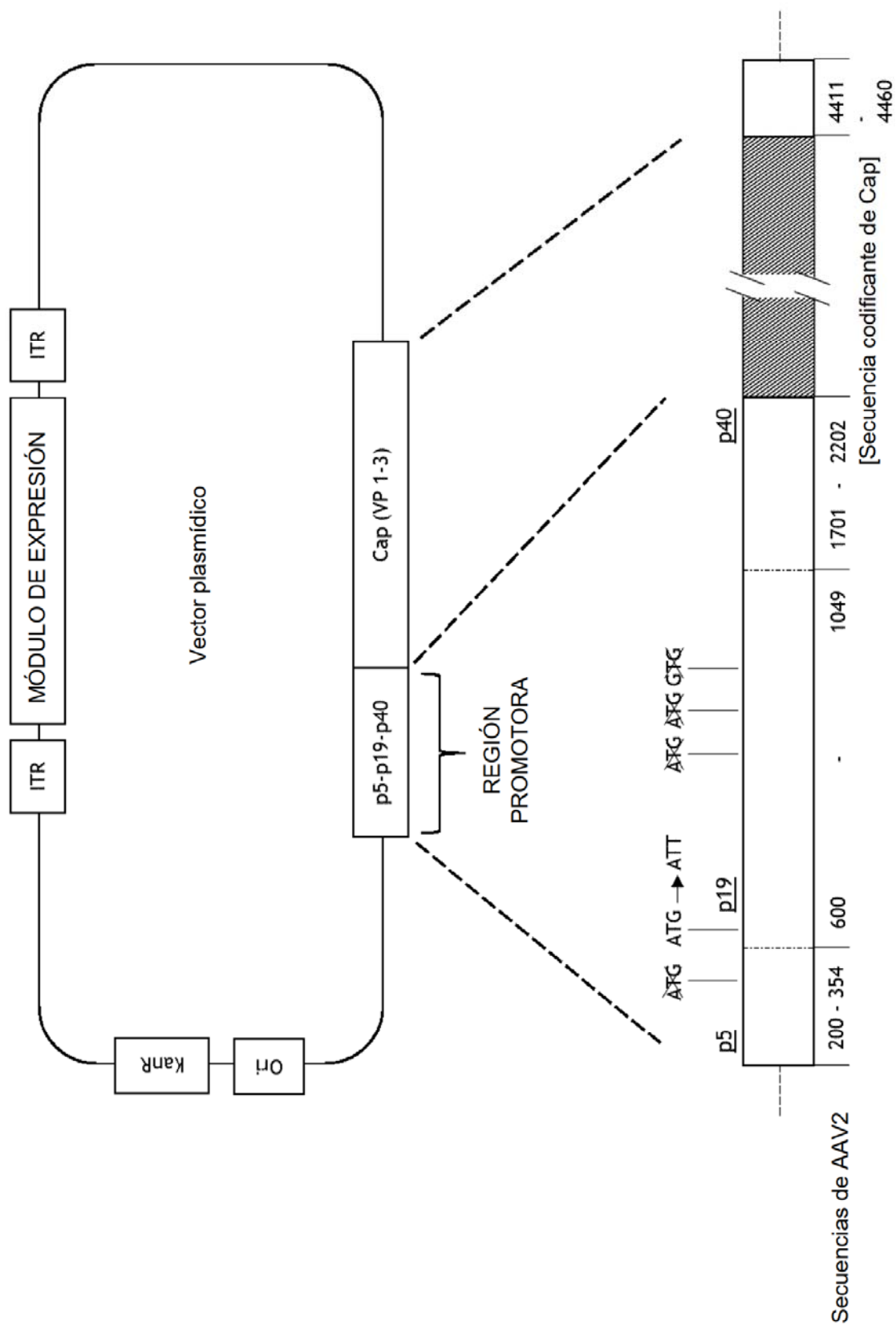


Fig. 4

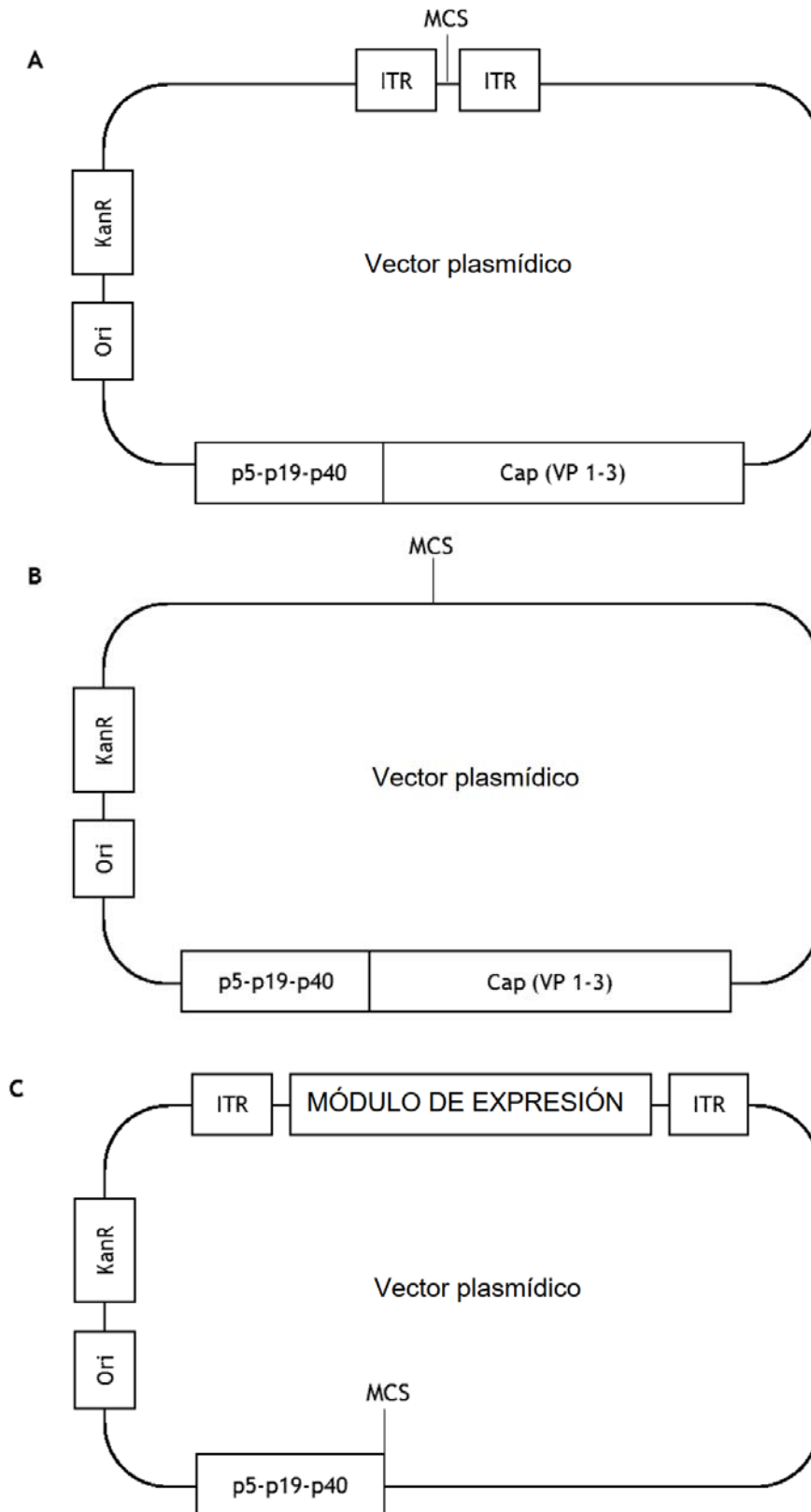


Fig. 4

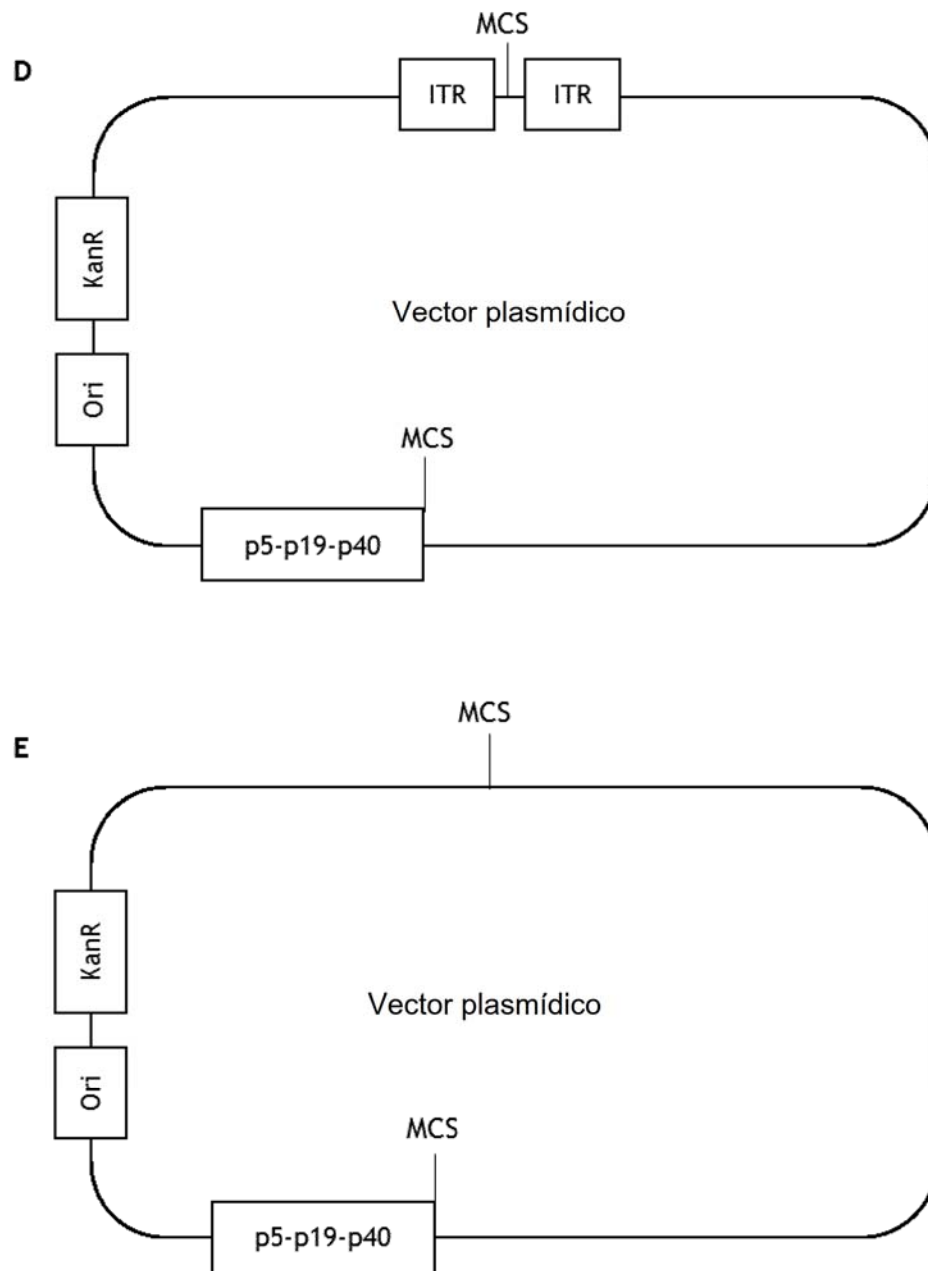
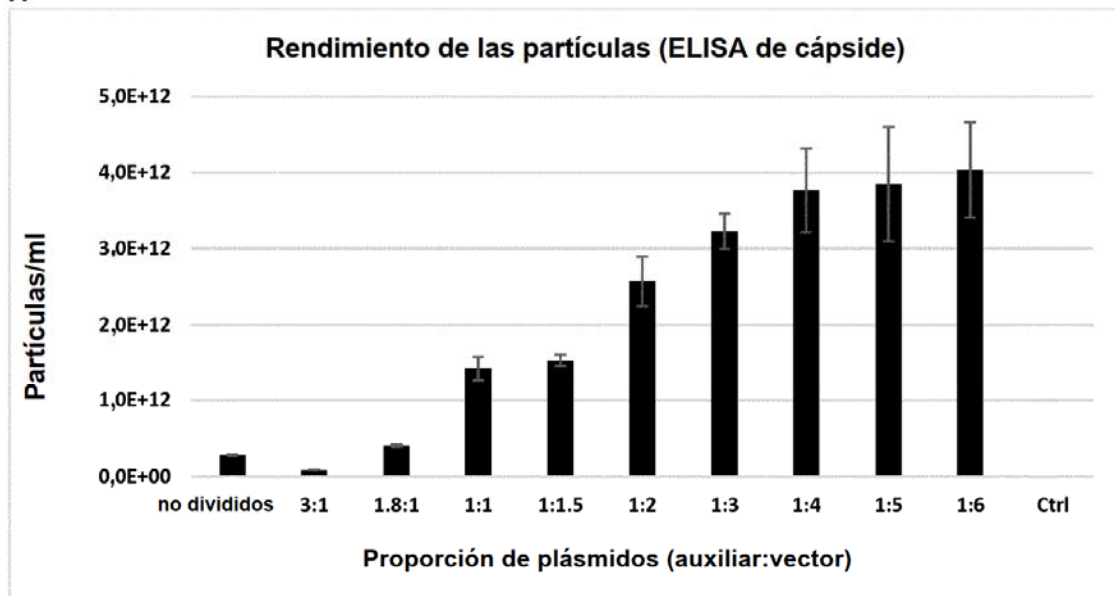


Fig. 5

A



B

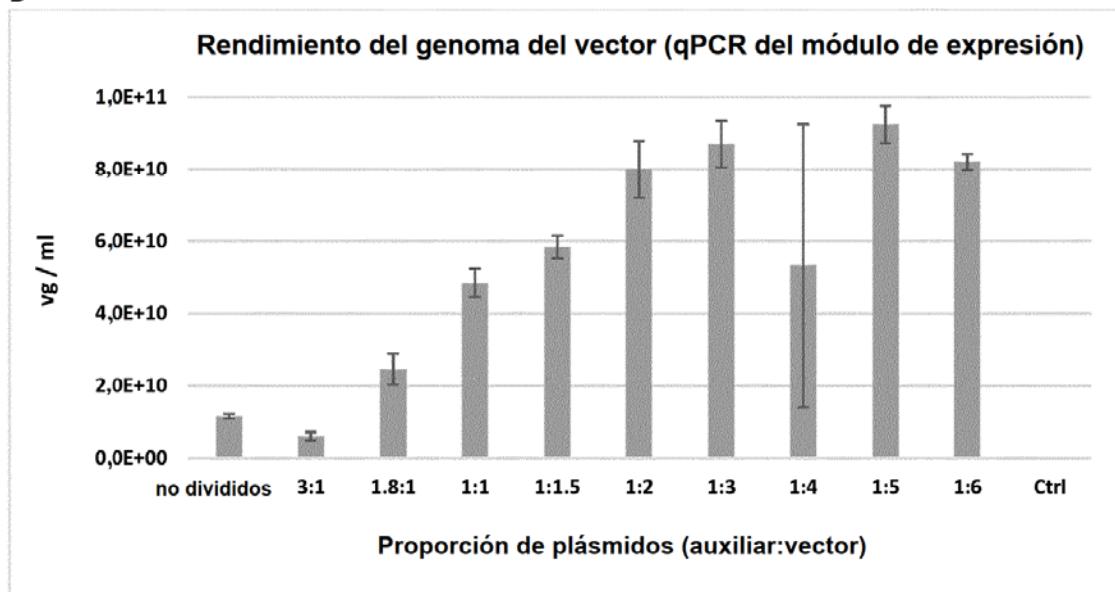


Fig. 5

C

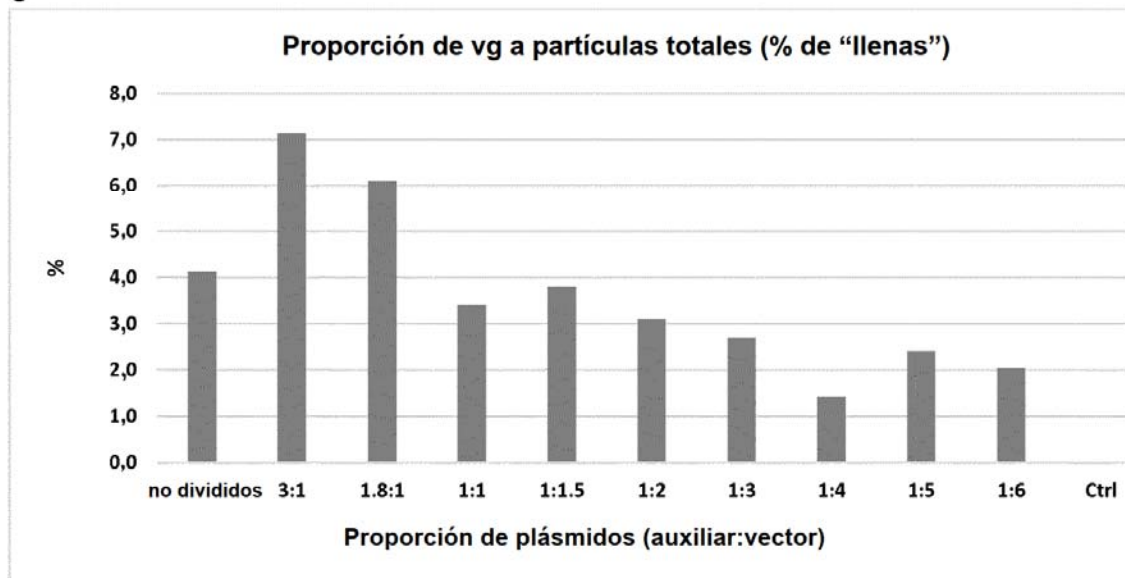


Fig. 6

	Resultado	Detección
Lote 1	Sin replicación	<10 rcAAV por 1E+10 vg
Lote 2	Sin replicación	<10 rcAAV por 1E+10 vg
Lote 3	Sin replicación	<10 rcAAV por 1E+10 vg

Fig. 7

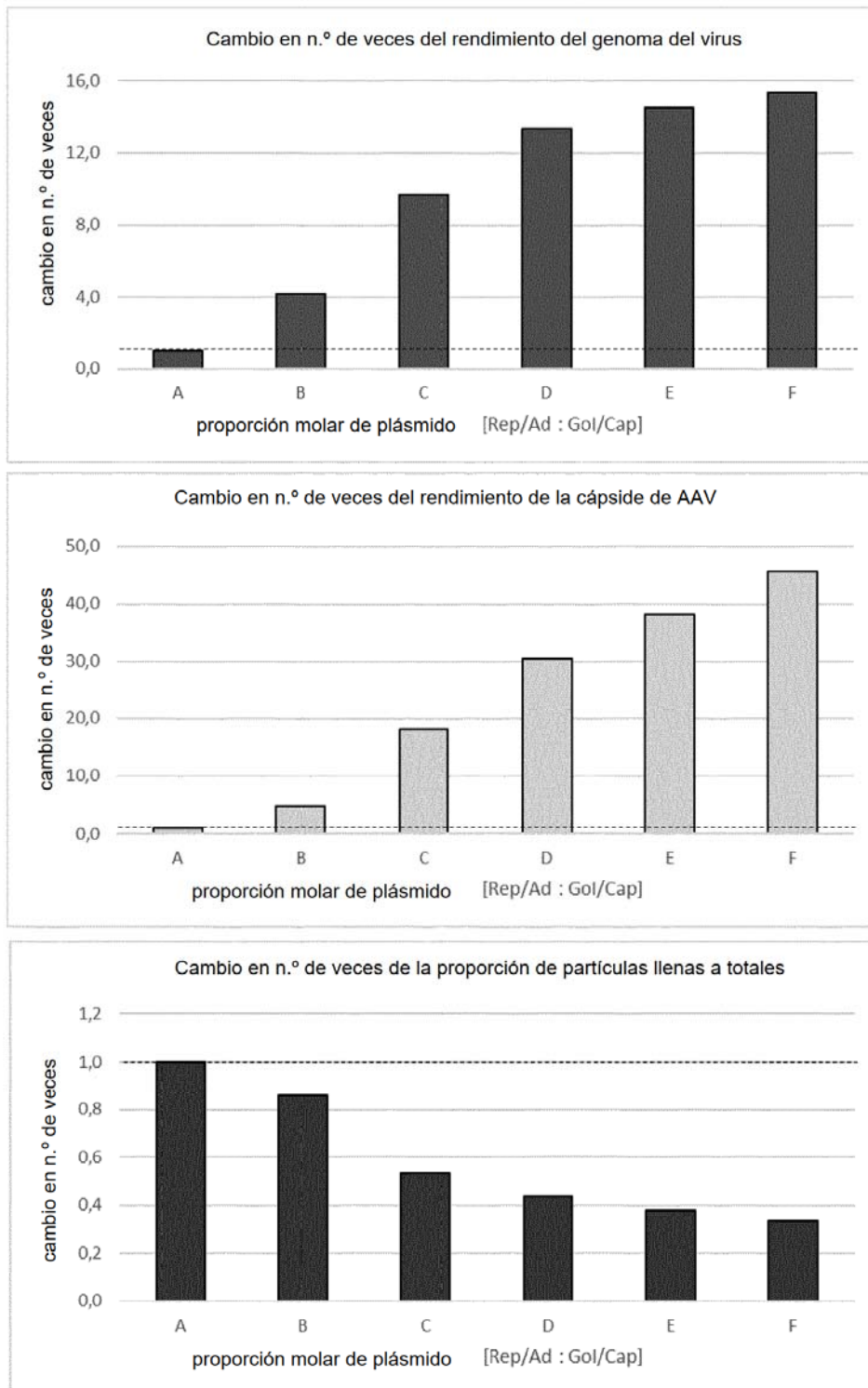


Fig. 8

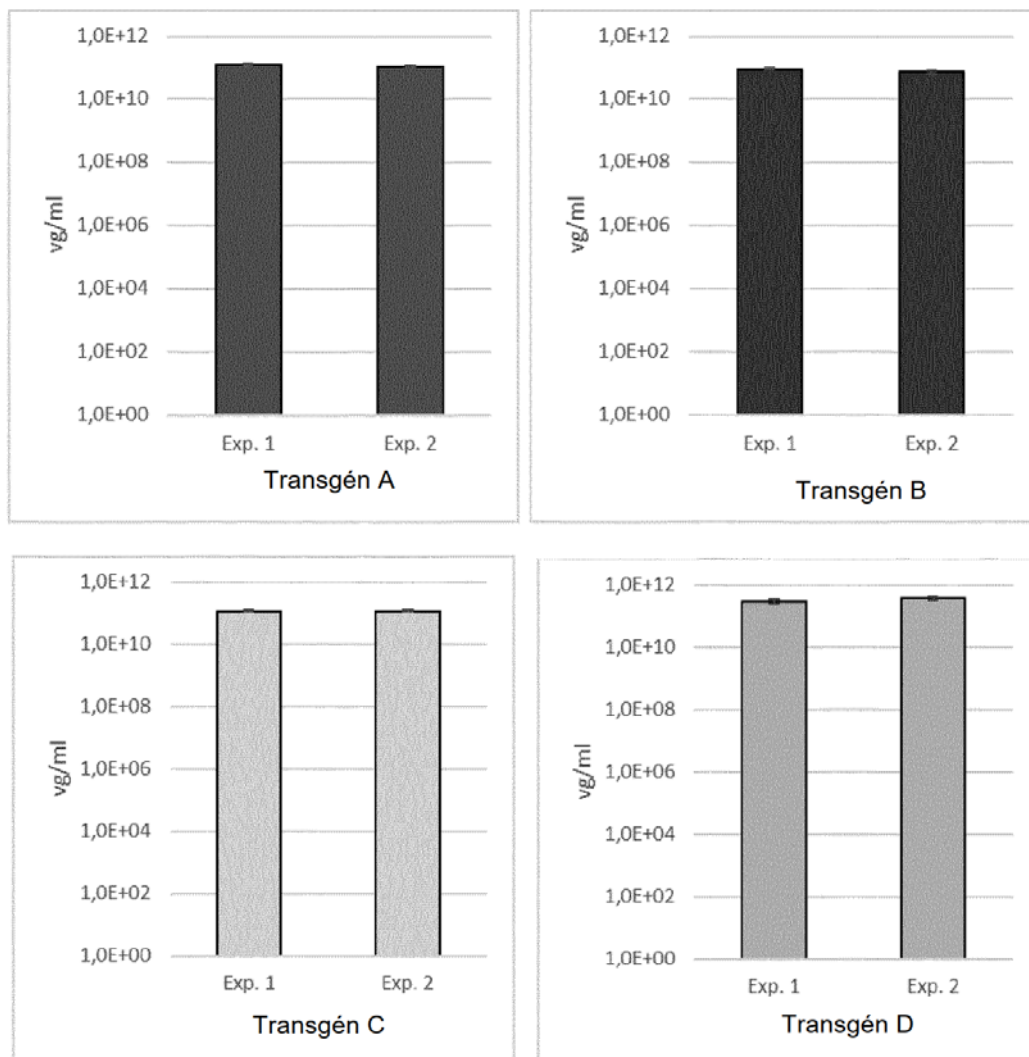


Fig. 9

