

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 661 447**

51 Int. Cl.:

C12N 15/52 (2006.01)

C12N 1/21 (2006.01)

C12P 19/62 (2006.01)

C07H 17/08 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.12.2012** **E 12198936 (2)**

97 Fecha y número de publicación de la concesión europea: **24.01.2018** **EP 2610344**

54 Título: **Identificación y caracterización del grupo de genes de biosíntesis de espinactina procedente de Saccharopolyspora spinosa productora de espinosina**

30 Prioridad:

28.12.2011 US 201161580947 P

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

02.04.2018

73 Titular/es:

DOW AGROSCIENCES LLC (100.0%)

9330 Zionsville Road

Indianapolis, Indiana 46268, US

72 Inventor/es:

PALANIAPPAN, NADARAJ;

LEWER, PAUL y

GRAUPNER, PAUL RICHARD

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 661 447 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

Identificación y caracterización del grupo de genes de biosíntesis de espinactina procedente de *Saccharopolyspora spinosa* productora de espinosina

Reivindicación de prioridad

- 5 Esta aplicación reivindica el beneficio de la Solicitud de Patente Provisional de los EE.UU. con nº de Serie 61/580.947, presentada el 28 de diciembre de 2011, titulada "IDENTIFICATION AND CHARACTERIZATION OF THE SPINACTIN BIOSYNTHESIS GENE CLUSTER FROM SPINOSYN-PRODUCING *SACCHAROPOLYSPORA SPINOSA*".

Campo técnico

- 10 La presente descripción se refiere de forma general a métodos para producir espinosina con bajo contenido de impurezas y a espinactina y su biosíntesis.

Antecedentes

- 15 Tal como se describe en la Patente de EE.UU. nº 5.362.634, el producto de fermentación A83543 es una familia de compuestos relacionados producidos por *Saccharopolyspora spinosa*. Los miembros conocidos de esta familia han sido referidos como factores o componentes, y a cada uno de ellos se le ha asignado una designación de letra identificadora. Estos compuestos se denominan en la presente memoria espinosina A, B, etc. Los compuestos de espinosina son útiles para controlar artrópodos, nematodos e insectos, en particular, de las especies Lepidoptera y Diptera, son bastante benignos para el medioambiente y presentan un perfil toxicológico atractivo.

- 20 Los compuestos de espinosina producidos de forma natural consisten en un sistema de anillos 5,6,5-tricíclico, condensado con una lactona macrocíclica de 12 miembros, un azúcar neutro (rhamnosa) y un amino azúcar (forosamina) (véase Kirst et al. (1991)). Si no está presente el amino azúcar, los compuestos han sido referidos como pseudoaglicona de A, D, etc., y si el azúcar neutro no está presente, entonces los compuestos han sido referidos como pseudoaglicona inversa de A, D, etc. Una nomenclatura más preferida es referirse a las pseudoagliconas como espinosina A 17-Psa, espinosina D 17-Psa, etc., y a las pseudoagliconas inversas como espinosina A 9-Psa, espinosina D 9-Psa, etc.

- 25 Los compuestos de espinosina producidos de forma natural pueden ser producidos mediante fermentación a partir de cultivos NRRL 18395, 18537, 18538, 18539, 18719, 18720, 18743 y 18823. Estos cultivos han sido depositados y separados de la colección de reservas de cultivo del "Midwest Area Northern Regional Research Center", del "Agricultural Research Service" del Departamento de Agricultura de los Estados Unidos, 1815 North University Street, Peoria, Ill., 61604. Pan et al (J. Bacteriology 2011, p. 3150-3151) describen el primer borrador de secuencia genómica del tipo de cepa de *Saccharopolyspora spinosa* NRRL 18395.

- 30 La Patente de EE.UU. nº 5.362.634 y la correspondiente Solicitud de Patente Europea nº 375316 A1 se refieren a las espinosinas A, B, C, D, E, F, G, H y J. Se dice que estos compuestos son producidos cultivando una cepa del nuevo microorganismo de *Saccharopolyspora spinosa* seleccionada entre NRRL 18395, NRRL 18537, NRRL 18538 y NRRL 18539.

- 35 La Patente WO 93/09126 se refiere a las espinosinas L, M, N, Q, R, S y T. En ella también se discuten dos cepas productoras de la espinosina J: NRRL 18719 y NRRL 18720, y una cepa que produce las espinosinas Q, R, S y T: NRRL 18823.

- 40 La Patente WO 94/20518 y la Patente de EE.UU. nº 5.670.486 se refieren a las espinosinas K, O, P, U, V, W e Y, y a derivados de las mismas. También se discute la cepa productora de espinosina K, NRRL 18743.

- La Patente WO 03/070908 describe una policétido sintasa de espinosina híbrida capaz de funcionar en *Saccharopolyspora spinosa* para producir una espinosina biológicamente activa. La policétido sintasa híbrida comprende un módulo de carga heterólogo asociado operativamente a una pluralidad de módulos extensores de *Saccharopolyspora spinosa*.

- 45 Uno de los retos para producir compuestos de espinosina surge del hecho de que se requiere un volumen de fermentación muy grande para producir una cantidad muy pequeña de espinosinas. Es altamente deseable incrementar la eficiencia de la producción de espinosinas y de este modo mejorar la disponibilidad de las espinosinas, a la vez que se reducen sus costes.

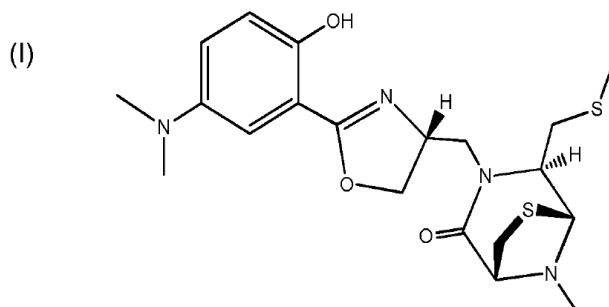
- 50 Otro reto es la producción de compuestos de espinosina mediante métodos que reduzcan o eliminen impurezas, y que no presenten efectos negativos sobre los niveles de producción de espinosinas.

Descripción

Un aspecto de la invención incluye un método para producir una cepa productora de espinosina de

Saccharopolyspora spinosa; método que comprende:

modificar una molécula de ácido nucleico que comprende una secuencia de nucleótidos que presenta una identidad de al menos 90% con respecto a SEQ ID NO: 1 mediante mutación de gen dentro de dicha secuencia de nucleótidos, de tal modo que el gen deja de expresar una proteína funcional; y de tal modo que la cepa mutada produce menos cantidad de un compuesto de Fórmula (I)



que la producida en la cepa antes de la mutación del gen.

Dicha secuencia modificada puede dar como resultado una molécula de ácido nucleico que codifica una espinosina que sintetiza un policétido diferente de los policétidos sintetizados por una cepa productora de espinosina nativa. En una realización, la secuencia de nucleótidos es SEQ ID NO: 2. En una realización, la cepa productora de espinosina produce espinosina J y L. En una realización, la mutación de la secuencia de nucleótidos comprende introducir una eliminación o un codón de parada.

Otro aspecto se refiere a una célula de la cepa productora de espinosina de *S. spinosa* obtenida mediante el método del aspecto anterior, en donde la célula comprende una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene al menos un 90% de identidad con respecto a SEQ ID NO: 1 y que comprende una mutación en un gen en dicha secuencia de nucleótidos, de tal forma que el gen deje de expresar una proteína funcional, y en donde la célula produce menos del compuesto de Fórmula (I) de lo que se producía en una célula de la misma cepa antes de la mutación del gen.

Otra realización incluye un método del aspecto anterior, que comprende:

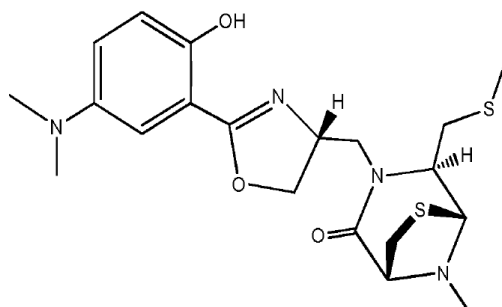
(i) proporcionar una molécula de ácido nucleico que comprende el gen mutado ligado operativamente a al menos una secuencia reguladora de la transcripción; en donde el gen es seleccionado del grupo que consiste en proteína de unión a sustrato de transportador ABC; permeasa de transportador ABC de aminoácido; transportador de resistencia a fármaco de la familia EmrB QacA; monooxigenasa; NRPS1; NRPS2; un gen localizado entre NRPS1 y NRPS2 en el genoma de *S. spinosa*; tiazolinil imida reductasa; tioesterasa; metiltransferasa; pabAB; un regulador transcripcional; Acil-CoA sintasa; N-metiltransferasa; piridoxamina 5'-fosfato oxidasa; 2,3-dihidrobenczoato-AMP ligasa; y aminotransferasa;

(ii) introducir la molécula de ácido nucleico en una población de células receptoras de tal modo que la secuencia de nucleótidos se integra de forma estable en el genoma; en donde la célula mantiene la producción de espinosina; y

(iii) cribar la célula receptora en función de una producción reducida del compuesto de Fórmula (I) que se producía en una célula de la misma cepa antes de la mutación del gen.

Determinadas realizaciones incluyen un método para producir una mezcla de espinosinas. El método incluye: proporcionar un cultivo de al menos un organismo de *Saccharopolyspora spinosa* que comprende un medio para alterar la biosíntesis de espinactina; cultivar al menos un organismo de *S. spinosa* en condiciones de fermentación; y obtener una mezcla de espinosinas a partir del cultivo de fermentación de *S. spinosa*. En otra realización, un método para producir espinosinas incluye: proporcionar un organismo de *Saccharopolyspora spinosa* modificado preparado según el método anteriormente mencionado; y cultivar el organismo en condiciones de fermentación, produciendo de este modo al menos una espinosina.

En la presente memoria también se describe un compuesto que tiene la fórmula (I):



(I),

o una sal o N-óxido del mismo. También se describe en la presente memoria un método para aislar el compuesto de fórmula (I) que incluye: cultivar una cepa de *Saccharopolyspora spinosa*, NRRL 18395, en condiciones de fermentación; obtener un material técnico que comprende metabolitos secundarios procedentes del cultivo de fermentación de *S. spinosa*; formar una suspensión del material técnico en una disolución que comprende metanol y agua en cantidades iguales (1:1); y filtrar la suspensión, de tal modo que se forme un filtrado que comprende una cantidad enriquecida del compuesto; y separar a continuación el compuesto (I) de los factores de espinosina mediante cromatografía líquida preparativa.

Las anteriores características, y otras, serán más evidentes a partir de la siguiente descripción detallada de varias realizaciones, que se realiza en referencia a las figuras acompañantes.

Breve descripción de las figuras

Figura 1: incluye una representación de la estructura química de la espinactina.

Figura 2: incluye el diagrama de modelación molecular "Oak Ridge Thermal Ellipsoid Plot Program" (ORTEP) correspondiente a espinactina determinado a partir de datos de cristalografía de rayos X.

Figura 3: incluye un mapa físico anotado del grupo de genes de espinactina putativo identificado a partir del contig 261 de la secuencia genómica de la cepa NRRL 18395 de *S. spinosa*.

Figura 4: incluye una ilustración de contig 261 que consiste en el grupo de genes de espinactina putativo de la secuencia genómica de *S. spinosa* que se muestra en relación a los tres clones cósmidos identificados con sus respectivos insertos. Se muestra la localización del fragmento parcial de *tiazolinil imida reductasa*; que sirvió como sonda en la estrategia de cribado.

Figura 5: incluye una ilustración del mapa de plásmido de *pCR81GWITopo_aac(3)-attB*.

Figura 6: incluye una ilustración de la eliminación de un fragmento de 7,6 kb del grupo de genes de espinactina en el clon de cósmido 4G15, que fue producida usando "Redirect Recombineering Technology™".

Listado de secuencias

Las secuencias de ácido nucleico incluidas en el listado de secuencias acompañante se muestran usando las abreviaturas estándar para las bases de nucleótidos, tal como se define en 37 C.F.R. § 1.822. Solo se muestra una cadena de cada secuencia de ácido nucleico, pero se entiende que la cadena complementaria está incluida en cualquier referencia a la cadena mostrada. En el listado de secuencias acompañante:

SEQ ID NO: 1, proporciona una secuencia de polinucleótido del contig 261 del genoma de *S. spinosa*;

SEQ ID NO: 2, proporciona una secuencia de polinucleótido del gen *NRPS1*;

SEQ ID NO: 3, proporciona una secuencia de polinucleótido de la secuencia de nucleótidos de 412 pb que contiene un fragmento de la secuencia codificadora de *tiazolinil imida reductasa*;

SEQ ID NO: 4, proporciona una secuencia de polinucleótido de un cebador directo de amplificación por PCR;

SEQ ID NO: 5, proporciona una secuencia de polinucleótido de un cebador inverso de amplificación por PCR;

SEQ ID NO: 6, proporciona una secuencia de polinucleótido de un fragmento que contiene el sitio de unión ϕ C31 de actinófono (*attB*) de *Streptomyces lividans*. Los sitios enzimáticos de restricción *Bam*HI y *Nde*I fueron añadidos en los extremos 5' y 3', respectivamente;

SEQ ID NO: 7, proporciona una secuencia de polinucleótido de cebador2 directo, un cebador de amplificación por PCR;

SEQ ID NO: 8, proporciona una secuencia de polinucleótido de cebador2 inverso, un cebador de amplificación por PCR;

- 5 SEQ ID NO: 9, proporciona una secuencia de polinucleótido de plásmido *pCR8/GWITPOO_aac(3)-attB*;

SEQ ID NO: 10, proporciona una secuencia de polinucleótido de cebador3 directo, un cebador de amplificación por PCR, las secuencias subrayadas se hibridan con la casete *aac(3)-attB*, las secuencias no subrayadas son del gen de *tiazolinil imida reductasa*;

TCGACGGCCTGCTGCGCCGCATCCTCAACCAGCCGACAATGTAGGCTG GAGCTGCTTC;

- 10 SEQ ID NO: 11, proporciona una secuencia de polinucleótido de cebador3 inverso, un cebador de amplificación por PCR, las secuencias subrayadas se hibridan con la casete *aac(3)-attB*, las secuencias no subrayadas son del gen de *tiazolinil imida reductasa*;

AACGCACCGCGCAGGCGCTGCGCGACGGATGGTTGAACACATATGGTC GACATGCCCGC; y

- 15 SEQ ID NO: 12, proporciona una secuencia de polinucleótido de unión donde el gen de *tiazolinil imida reductasa*; fue eliminado del clon cósmido 4G15, según se confirmó mediante secuenciación de ADN.

Modo(s) para llevar a cabo la invención

- 20 Los genes biosintéticos de espinosina y los ORFs relacionados fueron clonados y se determinó la secuencia de ADN de cada uno de ellos. Los genes y ORFs clonados se designan en la presente memoria como *spnA*, *spnB*, *spnC*, *spnD*, *spnE*, *spnF*, *spnG*, *spnH*, *spnI*, *spnJ*, *spnK*, *spnL*, *spnM*, *spnN*, *spnO*, *spnQ*, *spnR*, *spnS*, *ORFL15*, *ORFL16*, *ORFR1*, *ORFR2*, *S. spinosa gtt*, *S. spinosa gdh*, *S. spinosa epi*, y *S. spinosa kre*.

- 25 La *Saccharopolyspora spinosa* produce una mezcla de nueve compuestos relacionados estrechamente, denominados colectivamente "espinosinas". De esta mezcla, las espinosinas A y D, conocidas como spinosad, son los componentes principales y presentan la mayor actividad contra insectos diana claves. Las espinosinas J y L, dos de los componentes minoritarios de la mezcla de espinosinas, son los precursores del spinetoram, el insecticida de espinosina de segunda generación.

- 30 El spinosad es un insecticida producido por Dow AgroSciences (Indianapolis, Ind.) que está compuesto principalmente por aproximadamente un 85% de espinosina A y aproximadamente un 15% de espinosina D. Las espinosinas A y D son productos naturales producidos mediante fermentación de *Saccharopolyspora spinosa*, tal como se describe en la Patente de EE.UU. n° 5.362.634. Spinosad es un ingrediente activo de varias formulaciones de insecticidas disponibles comercialmente en Dow AgroSciences, que incluyen los productos para control de insectos TRACER™, SUCCESS™, SPINTOR™ y CONSERVE™. Por ejemplo, el producto TRACER está compuesto por entre aproximadamente 44% y aproximadamente 48% de Spinosad (p/v), o aproximadamente 4 libras (1,81 kg) de Spinosad por galón (3,78 litros). Los compuestos de espinosina en formulaciones granulares y líquidas han establecido su utilidad para el control de arácnidos, nematodos e insectos, en particular las especies Lepidoptera, Thysanoptera y Diptera. Las espinosinas A y D también se denominan en la presente memoria Espinosina A/D.

- 40 Spinetoram es una mezcla de 5,6-dihidro-3'-etoxi espinosina J (componente principal) y 3'-etoxi espinosina L producida por Dow AgroSciences. La mezcla se puede preparar etoxilando una mezcla de espinosina J y espinosina L, seguido de hidrogenación. El enlace doble 5,6 de la espinosina J y su 3'-etoxi es hidrogenado mucho más fácilmente que el de la espinosina L y su derivado 3'-etoxi, debido a impedimentos estéricos por el grupo metilo del C-5 en la espinosina L y su derivado 3'-etoxi. Véase la Patente de EE.UU. n° 6.001.981. Las espinosina J y L también se denominan en la presente memoria Espinosina J/L.

- 45 También se pueden producir nuevas espinosinas mediante mutagénesis de los genes clonados, y sustitución de los genes mutados por sus contrapartidas no mutadas en un organismo productor de espinosinas. La mutagénesis puede implicar, por ejemplo: 1) eliminación o desactivación de un dominio de cetorreductasa, deshidratasa o enoil reductasa (KR, DH o ER) de tal manera que una o más de dichas funciones es bloqueada y la cepa produce una espinosina que tiene un núcleo de lactona con un enlace doble, un grupo hidroxilo, o un grupo ceto que no está presente en el núcleo de la espinosina A (véase Donadio et al., 1993); 2) reemplazamiento de un dominio AT de tal modo que se incorpora un ácido carboxílico diferente en el núcleo lactona (véase Ruan et al., 1997); 3) adición de un dominio KR, DH o ER a un módulo PKS existente de tal modo que la cepa produce una espinosina que tiene un núcleo lactona con un enlace saturado, un grupo hidroxilo, o un doble enlace que no está presente en el núcleo de la espinosina A; o 4) adición o sustracción de un módulo de PKS completo de tal modo que el núcleo lactona cíclico tiene un número mayor o menor de átomos de carbono. Se puede crear un PKS híbrido reemplazando el dominio de carga de PKS de espinosina con carga de PKS heteróloga. Véase, p.ej., la Patente de EE.UU. n° 7.626.010.
- 55 También se ha destacado que las espinosinas vía modificación de los azúcares que están unidos a la cadena

principal de la lactona de espinosina pueden incluir modificaciones del resto rhamnosa y/o forosamina, o la unión de diferentes azúcares desoxi. El grupo de Salas en España demostró que se pueden producir nuevos compuestos de policétidos sustituyendo la molécula de azúcar existente por diferentes moléculas de azúcar. Rodríguez et al. J. Mol. Microbiol Biotechnol. 2000 Jul; 2(3): 271-6. Los ejemplos que se incluyen en la descripción ayudan a ilustrar el uso de la mutagénesis para producir una espinosina con funcionalidad modificada.

El ADN de la región de grupo de genes de espinosina se puede usar como sonda de hibridación para identificar secuencias homólogas. De esta manera, el ADN clonado aquí podría usarse para localizar plásmidos adicionales en bibliotecas génicas de *Saccharopolyspora spinosa* que solapen con la región descrita en la presente memoria, pero que también contengan ADN no clonado previamente de regiones adyacentes del genoma de la *Saccharopolyspora spinosa*. Adicionalmente, se puede usar ADN de la región clonada en la presente memoria para identificar secuencias no idénticas, aunque similares, en otros organismos. Las sondas de hibridación normalmente tienen una longitud de al menos 20 bases y están marcadas para permitir la detección.

Según una realización particular de la invención, un método para producir una cepa productora de espinosina comprende modificar una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene al menos un 90% de identidad con respecto a SEQ ID NO: 1 mediante mutación de un gen de dicha secuencia de nucleótidos, de tal forma que el gen deje de expresar una proteína funcional; y de tal forma que la cepa modificada produzca menos de un compuesto de Fórmula (I). Dicha secuencia mutada puede dar como resultado una molécula de ácido nucleico que codifica espinosina que sintetiza un policétido diferente al policétido sintetizado a partir de una cepa productora de espinosina nativa. La cepa productora de espinosina puede producir las espinosinas A y D, y en realizaciones particulares, puede producir las espinosinas J y L.

Los términos “condiciones estrictas” o “hibridación en condiciones estrictas” se refieren a las condiciones en las que una sonda se hibridará preferencialmente a su subsecuencia diana, y en menor medida, o nada en absoluto, a otras secuencias. “Hibridación estricta” y “condiciones de lavado de hibridación estricta” en el contexto de experimentos de hibridación de ácido nucleico tales como las hibridaciones Southern y Northern son dependientes de la secuencia, y son diferentes bajo parámetros ambientales diferentes. Un extensa guía sobre la hibridación de ácidos nucleicos se puede encontrar en Tijssen (1993) “*Laboratory Techniques in Biochemistry and Molecular Biology – Hybridization with Nucleic Acid Probes*” parte I, capítulo 2, “*Overview of principles of hybridization and the strategy of nucleic acid probe assays*”, Elsevier, Nueva York. Generalmente, se selecciona una hibridación altamente estricta y condiciones de lavado altamente estrictas para que sean aproximadamente 5°C inferiores al punto de fusión térmica (T_m) de la secuencia específica a una fuerza iónica y pH definidos. La T_m es la temperatura (para una fuerza iónica y un pH definidos) a la cual el 50% de la secuencia diana se hibrida con una sonda perfectamente coincidente. Las condiciones muy estrictas se seleccionan para que sean iguales a la T_m de una sonda particular.

Un ejemplo de condiciones de hibridación estrictas para la hibridación de ácidos nucleicos complementarios que tienen más de 100 residuos complementarios en un filtro en un ensayo Southern o Northern es 50% de formamida con 1 mg de heparina a 42°C, llevándose a cabo la hibridación durante una noche. Un ejemplo de condiciones de lavado altamente estrictas es NaCl 0,15 M a 72°C durante aproximadamente 15 minutos. Un ejemplo de condiciones de lavado estrictas es un lavado 0,2xSSC (tampón de cloruro sódico/citrato sódico) a 65°C durante 15 minutos (véase Sambrook et al. (1989) “*Molecular Cloning – A Laboratory Manual*” (2ª ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY, para una descripción de tampón SSC). A menudo, un lavado altamente estricto viene precedido de un lavado poco estricto para eliminar la señal de sonda de fondo. Un ejemplo lavado medio estricto para un dúplex de, p.ej., más de 100 nucleótidos, es 1xSSC a 45°C durante 15 minutos. Un ejemplo de lavado poco estricto para un dúplex de, p.ej., más de 100 nucleótidos, es 4-6xSSC a 40°C durante 15 minutos. En general, una ratio señal-ruido de 2x (o superior) respecto a la observada para una sonda no relacionada en el ensayo de hibridación particular indica la detección de una hibridación específica. Los ácidos nucleicos que no se hibridan unos a otros en condiciones de severidad siguen siendo sustancialmente idénticos si los polipéptidos que codifican son sustancialmente idénticos. Esto se produce, p.ej., cuando se crea una copia de un ácido nucleico usando la degeneración de codón máxima permitida por el código genético.

También se describe en la presente memoria un polinucleótido aislado hibridizable en condiciones estrictas, preferiblemente en condiciones altamente estrictas, a un polinucleótido descrito en relación a la presente invención.

Tal como se usa en al presente memoria, el término “hibridar” pretende describir las condiciones de hibridación y de lavado a las cuales secuencias de nucleótidos que son al menos aproximadamente un 50%, al menos aproximadamente un 60%, al menos aproximadamente un 70%, más preferiblemente al menos aproximadamente un 80%, incluso más preferiblemente al menos aproximadamente de un 85% a un 90%, lo más preferiblemente al menos un 95% homólogas entre ellas típicamente permanecen hibridadas unas a otras.

Otro ejemplo no limitante de condiciones de hibridación estrictas es la hibridación en 6x SSC a aproximadamente 45°C, seguida de uno o más lavados en 1xSSC, 0,1% de SDS a 50°C, preferiblemente a 55°C, más preferiblemente a 50°C e incluso más preferiblemente a 65°C.

Las condiciones altamente estrictas pueden incluir incubaciones a 42°C durante un periodo de varios días, tal como 2-4 días, usando una sonda de ADN marcada, tal como sonda de ADN marcada con digoxigenina (DIG), seguido de

uno o más lavados en 2xSSC, 0,1% de SDS a temperatura ambiente y uno o más lavados en 0,5xSSC, 0,1% de SDS o 0,1xSSC, 0,1% de SDS a 65-68°C. En particular, las condiciones altamente estrictas incluyen, por ejemplo, de 2h a 4 días de incubación a 42°C usando sonda de ADN marcada con DIG (preparada, p.ej., usando un sistema de marcaje DIG; Roche Diagnostics GmbH, 68298 Mannheim, Alemania) en una disolución tal como la disolución DigEasyHyb (Roche Diagnostics GmbH) con o sin 100 µg/mL de ADN de esperma de salmón, o una disolución que comprende un 50% de formamida, 5xSSC (NaCl 150 mM, citrato trisódico 15 mM), dodecil sulfato sódico al 0,02%, 0,1% de N-lauroilsarcosina y 2% de reactivo bloqueante (Roche Diagnostics GmbH), seguido de un lavado de los filtros dos veces durante entre 5 y 15 minutos en 2xSSC y 0,1% de SDS a temperatura ambiente, y después lavando dos veces durante 15-30 minutos en 0,5xSSC y 0,1% de SDS o 0,1xSSC y 0,1% de SDS a 65-68°C.

En algunas realizaciones una molécula de ácido nucleico aislado que se hibrida en condiciones altamente estrictas a una secuencia de nucleótidos de la invención puede corresponder a una molécula de ácido nucleico natural. Tal como se usa en la presente memoria, una molécula de ácido nucleico "natural" se refiere a una molécula de ARN o ADN que tiene una secuencia de nucleótidos que está presente en la naturaleza (p.ej., que codifica una proteína natural).

Un especialista sabrá qué condiciones aplicar para las condiciones de hibridación estrictas o altamente estrictas. En la técnica se encuentran fácilmente disponibles guías adicionales en relación a dichas condiciones, por ejemplo, en Sambrook et al., 1989, *"Molecular Cloning, A Laboratory Manual"*, Cold Spring Harbor Press, N.Y.; y Ausubel et al. (eds.), 1995, *"Current Protocols in Molecular Biology"*, (John Wiley & Sons, N.Y.).

"Polimorfismo funcional" tal como se usa en la presente memoria se refiere a un cambio en la secuencia de pares de bases de un gen que produce un cambio cualitativo o cuantitativo en la actividad de la proteína codificada por dicho gen (p.ej., un cambio en la especificidad de la actividad; un cambio en el nivel de actividad). El término "polimorfismo funcional" incluye mutaciones, eliminaciones e inserciones.

En general, la etapa de detección del polimorfismo de interés se puede llevar a cabo tomando una muestra biológica que contiene el ADN de la fuente, y a continuación determinando la presencia o la ausencia de ADN que contiene el polimorfismo de interés en la muestra biológica.

La determinación de la presencia o de la ausencia de ADN que codifica una mutación particular puede llevarse a cabo con una sonda de oligonucleótido marcada con un grupo detectable adecuado, y/o por medio de una reacción de amplificación tal como la reacción en cadena de polimerasa o la reacción en cadena de ligasa (el producto de dicha reacción de amplificación puede ser detectado a continuación con una sonda de oligonucleótido marcado o mediante una serie de técnicas adicionales). Además, la etapa de detección puede incluir la etapa de detectar si el sujeto es heterocigoto u homocigoto para la mutación particular. Se conocen numerosos formatos de ensayo de sonda de oligonucleótido diferentes que pueden emplearse para llevar a cabo la presente invención. Véase, p.ej., la Patente de EE.UU. n° 4.302.204 a nombre de Wahl et al.; la Patente de EE.UU. n° 4.358.535 a nombre de Falkow et al.; la Patente de EE.UU. n° 4.563.419 a nombre de Ranki et al.; y la Patente de EE.UU. n° 4.994.373 a nombre de Stavrianopoulos et al.

La amplificación de una secuencia de ácido nucleico diana seleccionada, o diana, se puede llevar a cabo mediante cualquier medio adecuado. Véase de forma general, Kwoh et al., Am. Biotechnol. Lab. 8, 14-25 (1990). Los ejemplos de técnicas de amplificación adecuadas incluyen, aunque sin limitación, reacción en cadena de polimerasa, reacción en cadena de ligasa, amplificación de desplazamiento de cadena (véase de forma general G. Walker et al., Proc. Natl. Acad. Sci. USA 89, 392-396 (1992); G. Walker et al., Nucleic Acids Res. 20, 1691-1696 (1992)), amplificación basada en transcripción (véase D. Kwoh et al., Proc. Natl. Acad. Sci. USA 86, 1173-1177 (1989)), replicación de secuencia auto-sostenida (o "3SR") (véase J. Guatelli et al., Proc. Natl. Acad. Sci. USA 87, 1874-1878 (1990)), el sistema de replicasa Q β (véase P. Lizardi et al., BioTechnology 6, 1197-1202 (1988)), amplificación basada en secuencia de ácido nucleico (o "NASBA") (véase R. Lewis, Genetic Engineering News 12 (9), 1 (1992)), la reacción en cadena de reparación (o "RCR") (véase R. Lewis, anterior), y la amplificación de ADN de boomerang (o "BDA") (véase R. Lewis, anterior). Generalmente se prefiere la reacción en cadena de polimerasa.

La reacción en cadena de polimerasa (PCR) se puede llevar a cabo de acuerdo a técnicas conocidas. Véanse, p.ej., las Patentes de EE.UU. n° 4.683.195; 4.683.202; 4.800.159 y 4.965.188. En general, la PCR implica, en primer lugar, el tratamiento de una muestra de ácido nucleico (p.ej., en presencia de una ADN polimerasa termoestable) con un oligonucleótido cebador para cada cadena de la secuencia específica que se va a detectar en condiciones de hibridación, de tal modo que se sintetiza un producto de extensión de cada cebador que es complementario a cada cadena de ácido nucleico, con los cebadores suficientemente complementarios a cada cadena de la secuencia específica para hibridarse con ella de tal modo que el producto de extensión sintetizado a partir de cada cebador, cuando es separado de su complemento, puede servir como plantilla para la síntesis del producto de extensión del otro cebador, y a continuación tratar la muestra en condiciones de desnaturalización para separar los productos de extensión del cebador de sus plantillas si la secuencia o secuencias a detectar están presentes. Estas etapas se repiten cíclicamente hasta que se obtiene el grado de amplificación deseado. La detección de la secuencia amplificada se puede llevar a cabo añadiendo al producto de reacción una sonda de oligonucleótido capaz de hibridarse al producto de reacción (p.ej., una sonda de oligonucleótido de la presente invención), portando la sonda una marca detectable, y a continuación detectando la marca según técnicas conocidas, o mediante visualización

directa sobre un gel. Dichas sondas pueden tener una longitud de entre 5 y 500 nucleótidos, preferiblemente entre 5 y 250, más preferiblemente entre 5 y 100 o entre 5 y 50 ácidos nucleicos. Cuando las condiciones de PCR permiten la amplificación de todos los tipos alélicos, los tipos pueden distinguirse mediante hibridación con una sonda específica alélica, mediante digestión con endonucleasa de restricción, mediante electroforesis o geles de gradiente desnaturalizante, o mediante otras técnicas.

La reacción en cadena de ligasa (LCR) también se lleva a cabo de acuerdo a técnicas conocidas. Véase, p.ej., R. Weiss, Science 254, 1292 (1991). En general, la reacción se lleva a cabo con dos pares de sondas de oligonucleótido: un par se une a una cadena de la secuencia a detectar; el otro par se une a la otra cadena de la secuencia a detectar. Cada par junto solapa completamente con la cadena a la que corresponde. La reacción se lleva a cabo, en primer lugar, desnaturalizando (p.ej., separando) las cadenas de la secuencia a detectar, a continuación haciendo reaccionar las cadenas con los dos pares de sondas de oligonucleótido en presencia de una ligasa termoestable de tal modo que cada par de sondas de oligonucleótido se ligan entre ellas, a continuación separando el producto de reacción, y después repitiendo cíclicamente el proceso hasta que la secuencia ha sido amplificada hasta el grado deseado. La detección se puede llevar a cabo entonces de un modo similar al descrito antes para la PCR.

Los términos "homología" o "porcentaje de identidad" se usan en la presente memoria de forma intercambiable. Para el propósito de esta invención, se define en la presente memoria que a fin de determinar el porcentaje de identidad de dos secuencias de aminoácido o de dos secuencias de ácido nucleico, las secuencias son alineadas con fines de comparación óptima (p.ej., se deben introducir huecos en la secuencia de una primera secuencia de aminoácidos o de ácido nucleico para un alineamiento óptimo con una segunda secuencia de aminoácidos o ácido nucleico). A continuación se comparan los residuos de aminoácido o los nucleótidos en las correspondientes posiciones de aminoácido o posiciones de nucleótido. Cuando una posición de la primera secuencia está ocupada por el mismo residuo de aminoácido o nucleótido que la posición correspondiente de la segunda secuencia, entonces las moléculas son idénticas en dicha posición. El porcentaje de identidad entre las dos secuencias es función del número de posiciones idénticas compartidas por las secuencias (es decir, % identidad = número de posiciones idénticas/número total de posiciones (es decir, posiciones solapadasx100). Preferiblemente, las dos secuencias tienen la misma longitud.

El especialista será consciente del hecho de que existen disponibles varios programas de ordenador diferentes para determinar la homología entre dos secuencias. Por ejemplo, una comparación de secuencias y la determinación del porcentaje de identidad entre dos secuencias se pueden realizar usando un algoritmo matemático. En una realización preferida, el porcentaje de identidad entre dos secuencias de aminoácido se determina usando el algoritmo de Needleman y Wunsch (J. Mol. Biol. (48): 444-453 (1970)) que se ha incorporado al programa GAP del paquete de software GCG (disponible en internet en la página web de accelrys, más específicamente en <http://www.accelrys.com>), usando una matriz Blossom 62 o una matriz PAM250, y un peso de hueco de 16, 14, 12, 10, 8, 6 ó 4, y un peso de longitud de 1, 2, 3, 4, 5 ó 6. El especialista apreciará que todos estos parámetros diferentes darán lugar a resultados ligeramente diferentes, pero el porcentaje de identidad global de las dos secuencias no se alterado significativamente cuando se emplean algoritmos diferentes.

En otra realización, el porcentaje de identidad entre dos secuencias de nucleótidos se determina usando el programa GAP del paquete de software GCG (disponible en internet en la página web de accelrys, más específicamente en <http://www.accelrys.com>), usando una matriz NWSgapdna.CMP y un peso de hueco de 40, 50, 60, 70 u 80 y un peso de longitud de 1, 2, 3, 4, 5 ó 6. En otra realización, el porcentaje de identidad entre dos secuencias de aminoácido o de nucleótidos se determina usando el algoritmo de E. Meyers y W. Miller (CABIOS, 4: 11-17 (1989), que ha sido incorporado al programa ALIGN (versión 2.0) (disponible en internet en la página web de vega, más específicamente ALIGN-IGH Montpellier, o más específicamente en <http://vega.igh.cnrs.fr/bin/align-guess.cgi>) usando una tabla de residuos de peso PAM120, un penalización de longitud de hueco de 12 y una penalización de hueco de 4.

Las secuencias de ácido nucleico y de proteína descritas en la presente memoria en relación a la presente invención pueden usarse adicionalmente como "secuencia de consulta" para llevar a cabo una búsqueda en bases de datos públicas para, por ejemplo, identificar otros miembros de la familia o secuencias relacionadas. Dichas búsquedas se pueden llevar a cabo usando los programas BLASTN y BLASTX (versión 2.0) de Altschul et al. (1990) J. Mol. Biol. 215: 403-10. Las búsquedas de nucleótidos con BLAST se pueden llevar a cabo con el programa BLASTN, puntuación=100, longitud de palabra=12, para obtener secuencias de nucleótidos homólogas a las moléculas de ácido nucleico de la presente invención. Se pueden llevar a cabo búsquedas de proteínas BLAST con el programa BLASTX, puntuación=50, longitud de palabra=3, para obtener secuencias de aminoácidos homólogas a las moléculas de proteína de la presente invención. Para obtener alineamientos con huecos con fines comparativos, se puede utilizar Gapped BLAST tal como se describe en Altschul et al., (1997) Nucleic Acids Res. 25 (17): 3389-3402. Cuando se utilizan los programas BLAST y Gapped BLAST, se pueden usar los parámetros por defecto de los respectivos programas (p.ej., BLASTX y BLASTN). (Disponible en internet en la página web de ncbi, más específicamente en <http://www.ncbi.nlm.nih.gov>).

Otra realización puede incluir una célula de la cepa de *S. spinosa* productora de espinosinas obtenida mediante el método de la presente invención mencionado anteriormente, donde la célula comprende una molécula de ácido

nucleico que comprende una secuencia de nucleótidos que tiene al menos un 90% de identidad con respecto a SEQ ID NO: 1 y que comprende una mutación en un gen dentro de dicha secuencia de nucleótido, de tal modo que el gen ya no expresa una proteína funcional, y en donde la célula produce menos del compuesto de Fórmula (I) de lo que se produce en una célula de la misma cepa antes de la mutación del gen. Un método alternativo también puede producir un policétido modificado que provoque la expresión de una molécula de ácido nucleico en una célula hospedante, policétido que también puede ser purificado adicionalmente. En una realización particular, la modificación de una secuencia puede comprender mutar, eliminar, reemplazar o inactivar una secuencia de las siguientes maneras: proporcionando una biblioteca de ácidos nucleicos, ácidos nucleicos que comprenden uno o más segmentos de polinucleótido ligados operativamente a al menos una secuencia reguladora de la transcripción; introduciendo la biblioteca de ácidos nucleicos en una población de células receptoras o de orgánulos intracelulares receptores, mediante lo cual dichos ácidos nucleicos desmontan la producción de espinactina a la vez que mantienen la producción de espinosina; e identificando al menos una célula receptora, orgánulo intracelular receptor u organismo que comprende una célula receptora, con un fenotipo deseado, controlando de este modo la cepa productora de espinosina. La mutación de la secuencia de nucleótidos puede incluir la introducción de una eliminación, una mutación o un codón de parada. La secuencia de nucleótidos puede estar comprendida en un gen que puede ser seleccionado del grupo que consiste en: proteína de unión a sustrato de transportador ABC; permeasa de transportador ABC de aminoácido; transportador de resistencia a fármaco de la familia EmrB QacA; monooxigenasa; NRPS1; NRPS2; un gen localizado entre NRPS1 y NRPS2 en el genoma de *S. spinosa*; tiazolinil imida reductasa; tioesterasa; metiltransferasa; pabAB; un regulador transcripcional; Acil-CoA sintasa; N-metiltransferasa; piridoxamina 5'-fosfato oxidasa; 2,3-dihidrobenczoato-AMP ligasa; y aminotransferasa. La secuencia de nucleótidos del gen de *tiazolinil imida transferasa*; puede ser al menos un 80% idéntica a SEQ ID NO: 3.

El organismo de *S. spinosa* modificado puede producir menos espinactina de la producida en el organismo de *S. spinosa* hospedante antes de la introducción del ácido nucleico, o puede no producir espinactina. En algunas realizaciones, el organismo de *S. spinosa* modificado puede producir: una cantidad de espinosina A que es sustancialmente idéntica a la cantidad de espinosina A producida en el organismo de *S. spinosa* hospedante antes de la introducción del ácido nucleico; una cantidad de espinosina D que es sustancialmente idéntica a la cantidad de espinosina D producida en el organismo de *S. spinosa* hospedante antes de la introducción del ácido nucleico, y una cantidad de espinactina que es inferior a la cantidad de espinactina producida en el organismo de *S. spinosa* hospedante antes de la introducción del ácido nucleico.

También se proporciona un organismo según el aspecto mencionado anteriormente en donde el organismo de *S. spinosa* modificado produce:

una cantidad de espinosina A que es sustancialmente idéntica a la cantidad de espinosina A producida en el organismo de *S. spinosa* hospedante antes de la introducción del ácido nucleico,

y

una cantidad de espinosina D que es sustancialmente idéntica a la cantidad de espinosina D producida en el organismo de *S. spinosa* hospedante antes de la introducción del ácido nucleico.

También se proporciona un método para producir espinosinas, método que comprende: cultivar el organismo de *Saccharopolyspora spinosa* modificado del aspecto mencionado anteriormente en condiciones de fermentación, produciendo así al menos una espinosina.

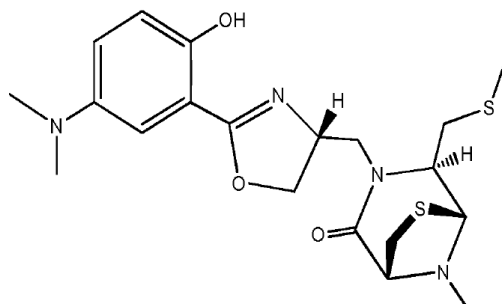
En una realización, el método comprende la obtención de una mezcla de espinosinas del cultivo de fermentación de *S. spinosa*. La alteración de la biosíntesis de espinactina se puede lograr como se ha descrito en los aspectos mencionados anteriormente. La mezcla de espinosinas obtenida del cultivo de fermentación de *S. spinosa* puede incluir al menos una de espinosina A y espinosina D. La mezcla de espinosinas obtenida del cultivo de fermentación de *S. spinosa* puede no incluir nada de espinactina. En algunas realizaciones, el gen del grupo de genes de espinactina es un gen de *tiazolinil imida reductasa*. Se puede producir un organismo de *Saccharopolyspora spinosa* modificado mediante el método. El organismo puede incluir un transgén implicado en la biosíntesis de espinosina. El organismo de *S. spinosa* modificado puede incluir un transgén implicado en la biosíntesis de espinosina que está integrado de forma estable en el ADN genómico del organismo en un sitio en donde el ácido nucleico fue integrado de forma estable.

Otra realización de la invención incluye un organismo de *Saccharopolyspora spinosa* modificado como se ha descrito en el aspecto mencionado anteriormente, que comprende un medio para alterar la síntesis de espinactina, en donde el organismo es capaz de producir espinosina A y/o espinosina D sin producir una cantidad significativa de espinactina.

También se describe en la presente memoria un método para producir un organismo de *Saccharopolyspora spinosa* modificado. El método incluye: proporcionar un ácido nucleico que comprende una secuencia de nucleótidos en el grupo de genes de biosíntesis de espinactina de la *S. spinosa*; mutar la secuencia de nucleótidos; e introducir el ácido nucleico que comprende la secuencia de nucleótidos en un organismo de *S. spinosa* hospedante, de tal manera que el ácido nucleico se integra de forma estable en el ADN genómico del organismo de *S. spinosa* hospedante, produciendo con ello un organismo de *S. spinosa* modificado, en donde un gen del grupo de genes de biosíntesis de espinactina ya no expresa una proteína funcional.

En una realización adicional, un método de la presente invención para producir espinosinas incluye: proporcionar un organismo de *Saccharopolyspora spinosa* preparado según el método mencionado anteriormente; y cultivar el organismo en condiciones de fermentación, produciendo de este modo al menos una espinosina.

En la presente memoria también se describe un compuesto que tiene la fórmula (I):



(I),

o una sal o N-óxido del mismo. También se describe en la presente memoria un método para aislar el compuesto de fórmula (I) que incluye: cultivar una cepa de *Saccharopolyspora spinosa*, NRRL 18395, en condiciones de fermentación; obtener un material técnico que comprende metabolitos secundarios procedentes del cultivo de fermentación de *S. spinosa*; formar una suspensión del material técnico en una disolución que comprende metanol y agua en cantidades iguales (1:1); y filtrar la suspensión, de tal modo que se forme un filtrado que comprende una cantidad enriquecida del compuesto.

Los siguientes ejemplos se proporcionan para ilustrar determinadas características y/o realizaciones particulares.

EJEMPLOS

Ejemplo 1: Identificación y caracterización de la molécula de espinactina procedente de *S. spinosa*

La espinactina es un compuesto que contiene azufre que se identificó en un cultivo de fermentación de *S. spinosa*. La Figura 1 ilustra la estructura química de la espinactina. Esta estructura se determinó usando espectroscopía de RMN y se verificó mediante cristalografía de rayos X. Véase la Figura 2.

Un cultivo de fermentación de *S. spinosa* de la cepa NRRL 18395 fue propagado usando el protocolo de fermentación descrito en Strobel y Nakatsukasa (1993) J. Ind. Microbiol. Biotechnol. 11(2): 121-7. Un análisis de cromatografía de líquidos identificó la presencia de un metabolito secundario no detectado previamente que presenta un espectro UV característico, con $\lambda_{\text{max}} = 227$ y 358 nm. El metabolito secundario se aisló usando el siguiente procedimiento. El material técnico (3 kg) procedente de un cultivo de fermentación de *S. spinosa* se puso en una suspensión en metanol/agua (1:1) y se filtró. Se observó que el filtrado contenía una cantidad enriquecida del metabolito secundario no detectado previamente. Dicho metabolito secundario fue separado del material de espinosina remanente mediante cromatografía líquida de fase inversa preparativa. Usando este método, se obtuvieron 6 g de la nueva molécula, que contabilizó aproximadamente el 0,02% en masa del material técnico total. La nueva molécula, que se denominó espinactina, fue re-cristalizada en metanol/agua en forma de un sólido verde claro (p.f. 79°C). La espectrometría de masas de impacto electrónico de sonda de alta resolución (EI/MS) indicó que la espinactina poseía la fórmula molecular de $\text{C}_{20}\text{H}_{28}\text{N}_4\text{O}_3\text{S}_2$. El análisis de los espectros de ^1H , ^{13}C y de aumento sin distorsión por transferencia de polarización (DEPT) indicaron la presencia de un grupo de dimetilamina, una segunda metilamina y un único grupo S-metilo. También fue evidente un anillo de benceno 1,2,4-tri-sustituido. El resto del espectro en cloroformo incluía multipletes complejos no de primer orden, que incluían efectos de acoplamiento virtual observados en espectros 2-D, producidos por los desplazamientos químicos coincidentes para 4-H y 5-H. De este modo, la estructura fue resuelta en metanol- d_4 , en donde el espectro fue de primer orden como se describe a continuación usando técnicas de RMN 2-D que incluyen COSY, HSQC y HMBC (ver la Tabla 1).

Se identificaron fácilmente tres sistemas de espín en el espectro de ^1H ; se detectaron tres multipletes aromáticos entre 7,3 y 6,7 ppm, que indican un anillo de benceno 1,2,4-tri-sustituido que contiene sustituyentes de oxígeno o nitrógeno; otro sistema de espín de dos metilenos (un sustituido con oxígeno, y uno con nitrógeno) con una metina entre ambos, que también indicó una sustitución de nitrógeno; y un sistema de espín elongado que contenía tres metinas contiguas, con metilenos en los extremos. Ninguno de los multipletes resonó en un campo más elevado de 2,5 ppm, lo que indica un alto grado de sustitución de heteroátomos. Aparte de estos sistemas de espín, había tres singletes de metilo, que se demostró que eran un grupo dimetil amino, y otro grupo N-metilo, y un metil tioéter. La constante de acoplamiento entre dos de las metinas fue pequeña, lo que sugiere que los átomos de carbono no estaban conectados directamente. Estos dos átomos de carbono presentaron picos cruzados HMBC procedentes del grupo N-metilo, lo que indica que estaban a cualquiera de los lados de un nitrógeno terciario metilado.

La molécula se construyó mediante un análisis extensivo de los experimentos HMBC. De esta manera, se identificó

un anillo de oxazolina con unión al anillo aromático en C-2. Se determinó que este anillo presentaba un grupo metileno unido a C-5, que se usó como palanca para ayudar en el montaje del sistema de anillo de diazabicilooctano. Este metileno (que resuena a 3,9 y 3,5 ppm) indicó picos cruzados con una de las metinas del sistema de espín extendido (63 ppm), y también a un carbonilo de amida a 172 ppm. Otros picos cruzados con este carbonilo procedían de otra metionina (4,1 ppm) y uno de los otros grupos metileno (que resuena a 3,5 y 3,2 ppm) indicando un anillo de 6 miembros que contiene 2 átomos de nitrógeno. El hecho de este metileno estaban enlazado de vuelta con el anillo, fue determinado a través de la existencia de picos cruzados con ambas de las metinas conectadas con el átomo de nitrógeno metilado, el tioéter deducido por desplazamiento químico. El montaje del resto de la molécula se obtuvo fácilmente una vez determinado el sistema del anillo.

10 **Tabla 1.** Espectro de RMN de espinactina en MeOH.

Nº	¹ H	¹³ C	Nº	¹ H	¹³ C
2	-	167,8	14	3,15, dd (12,9, 4,3)	35,9
4	4,54, dd (8,7, 9,6)	71,3		2,79, dd (12,9, 4,3)	
	4,22, t (8,1)		15	2,15, s	16,0
5	4,65, m	65,8	16	2,40, s	40,0
6	3,86, dd (14,4, 7,0)	46,6	1'	-	154,0
	3,53, dd (14,4, 4,5)		2'	-	111,4
8	4,02, ddd (10,6, 4,3, 4,3)	62,7	3'	7,14, d (3,1)	114,1
9	5,11, dd (3,5, 1,6)	78,3	4'	-	145,7
11	4,09, dd (5,6, 1,3)	71,5	5'	7,08, dd (9,0, 3,1)	122,9
12	-	172,7	6'	6,90, d (9,0)	118,2
13	3,22, dd (5,8, 10,7)	34,5	NMe ₂	2,86, s	42,6
	3,53, d (10,7)				

Se usó un método de cristalografía de rayos X para validar la estructura de la molécula de espinactina. En la Figura 2 se muestra un modelo ORTEP (Oak Ridge Thermal Ellipsoid Program) de la estructura. Se irradió un cristal individual de espinactina (PF 79°C) de dimensiones 0,50 x 0,8 x 0,06 mm con radiación Mo K_α en un difractómetro de rayos X Nonius Kappa CCD (Bruker-Nonius, Madison, WI). Se recogieron un total de 7.801 reflexiones (de las cuales 2.276 eran únicas). Los datos se refinaron y se resolvieron usando métodos estándar y utilizando programas de análisis de datos SHELXS97 y SIR97.

Ejemplo 2: Identificación y caracterización del grupo de genes de biosíntesis de espinactina.

En base a la semejanza estructural del resto 2-hidroxifeniloxazolina de la espinactina en comparación el resto 2-hidroxifeniloxazolina de la micobactina A, se planteó la hipótesis de que ambas moléculas son producidas a través de rutas biosintéticas similares. La implicación del grupo de genes de biosíntesis de micobactina A y del gen *feniloxazolina sintasa mbtB* (*mbtB*) en la biosíntesis del resto de 2-hidroxifeniloxazolina de la micobactina A ha sido caracterizada en Quadri et al. (1998) Chemistry & Biology 15: 631-45. La secuencia de aminoácidos del *feniloxazolina sintasa mbtB* (Número de acceso: ZP_06505539) del grupo de genes de micobactina de la cepa 02_1987 de *Mycobacterium tuberculosis* se usó para cribar, *in silico*, la secuencia genómica obtenida de la cepa natural de *S. spinosa* (NRRL 18395). La búsqueda BLAST resultante identificó la presencia de una secuencia génica putativa procedente del genoma de *S. spinosa* que codificaba una secuencia de aminoácidos con similitud con MbtB. El gen *NRPS1* del genoma *S. spinosa* compartía aproximadamente un 46% de identidad a nivel de aminoácido con el gen *mbtB* del grupo de genes de *M. tuberculosis*. La identificación de este gen condujo a la posterior identificación de un grupo de genes de *S. spinosa* que contenía genes adicionales de los que se propone su implicación en la biosíntesis de espinactina. Estas secuencias génicas precedentes del grupo de genes de espinactina fueron anotadas usando el programa de ordenador, FgenesB (Universidad de Cambridge, R.U.). La Figura 3 muestra una anotación de la secuencia de polinucleótidos para el contig 261, que contiene la secuencia del gen *NRPS1* procedente de la secuencia genómica de *S. spinosa*, la secuencia de polinucleótido correspondiente al contig 261 se presenta como SEQ ID NO: 1. La secuencia de polinucleótido de *NRPS1* se presenta como SEQ ID NO: 2.

Ejemplo 3: Cribado e identificación de clones de cósmido que contienen el grupo de genes de espinactina procedente de una biblioteca de cósmidos genómica de *S. spinosa*

Una biblioteca de cósmidos genómica de *S. spinosa* (Solicitud de Patente de EE.UU. n° 13/100202) se sometió a cribado para identificar clones de cósmido que contenían el gen de *tiazolinil imida reductasa*. El gen *tiazolinil imida reductasa*; está localizado por debajo del gen *NRPS1*, y se caracterizó como un miembro putativo del grupo de genes de espinactina que codifica una enzima que no es esencial para la biosíntesis de espinosina A/D. Este gen fue usado para identificar clones de cósmido que podrían ser modificado eliminando un fragmento de 7,6 kb que contiene el grupo de genes de espinactina. Los clones de cósmido resultantes fueron usados a continuación para generar cepas de *S. spinosa* bloqueadas.

Un fragmento de 412 pb del gen de *tiazolinil imida reductasa*; fue preparado con un kit de marcaje de digoxigenina (DIG) de PCR. El fragmento de 412 pb (SEQ ID NO: 3) fue amplificado con PCR usando un cebador director (SEQ ID NO: 4) y un cebador inverso (SEQ ID NO: 5). La hibridación se completó a 42°C en tampón de hibridación suministrado comercialmente (Roche) durante una noche. Las membranas de nylon fueron lavadas en condiciones estrictas durante 5 minutos, dos veces en 2x SSC, 0,1% de SDS a temperatura ambiente y durante 15 minutos, dos veces en 0,1 x SSC, 0,1% de SDS a 68°C. El marcaje quimioluminiscente y la detección se llevaron a cabo usando el kit "*DIG Luminescent Detection Kit™*" (Roche). Se identificaron tres clones de cósmido positivos (2C14, 3E11 y 4G15) y se confirmaron mediante secuenciamiento final del inserto cósmido usando cebadores de secuenciamiento T3 y T7. La Figura 4 resume los detalles de los tres clones de cósmido identificados y sus tamaños de inserto de fragmento respectivos y la localización correspondiente al contig 261 y al fragmento de 412 pb del gen de *tiazolinil imida reductasa*; de la secuencia genómica de *S. spinosa*.

Ejemplo 4: Construcción de una casete de gen de resistencia a apramicina con un sitio de unión attB para la eliminación parcial del grupo de genes de espinactina

La construcción de una casete de apramicina-attB se completó clonando la secuencia de *attB* en una casete de gen de resistencia a apramicina. La síntesis de la secuencia del sitio de unión de actinófago ϕ C31 de *Streptomyces lividans* (*attB*; N° de acceso: X60952) se llevó a cabo por parte de un suministrador de servicios de una tercera parte (Integrated DNA Technologies, Coralville, IA). Para facilitar la clonación del fragmento de attB sintetizado, se añadieron sitios *Bam*H1 y *Nde*I a los extremos 5' y 3', respectivamente (SEQ ID NO: 6). El vector que contenía la secuencia de *attB* sintetizada se marcó como pIDTSMART.

El plásmido plantilla, pIJ773 (N° de acceso: AX657063), que contiene la casete de gen de resistencia de apramicina, *aac(3)-IV* (N° de acceso: X99313), y el *oriT* del plásmido RP4 (N° de acceso: L27758), flanqueado por sitios *FRT* se amplificó con PCR usando un cebador2 directo (SEQ ID NO: 7) y un cebador2 inverso (SEQ ID NO: 8). El fragmento de PCR se clonó en el vector pCR8/GW/TOPO™ (Invitrogen, Carlsbad, CA) siguiendo las instrucciones del protocolo del fabricante, y se confirmó la presencia y la integridad de la casete de apramicina insertada mediante digestión de restricción y secuenciamiento.

El fragmento *attB* de pIDTSMART fue digerido con *Bam*HI y *Nde*I y el fragmento de ADN de 300 pb resultante fue purificado en gel usando el kit "*QIAquick Gel Extraction Kit™*" (Qiagen, Valencia, CA). El fragmento *attB* purificado fue clonado en el vector pCRB/GW/TOPO™, que contenía la casete de gen de resistencia a apramicina en los sitios de restricción *Nde*I y *Bam*HI, usando técnicas estándar de biología molecular. La construcción de una nueva casete de gen de resistencia a apramicina que contiene un sitio de unión a *attB* se confirmó mediante digestión con enzimas de restricción y secuenciamiento. El mapa y la secuencia del plásmido construido *pCR8/GW/TOPO_aac(3)-attB* se proporcionan en la Figura 5 (SEQ ID NO: 9).

Ejemplo 5: Generación de cepas de *S. spinosa* bloqueadas que contienen una eliminación parcial del grupo de genes de espinactina

Se obtuvo una eliminación de 7,6 Kb del grupo de genes de espinactina atacando a la región de *tiazolinil imida reductasa*; con una casete de *apramicina-attB* usando la tecnología "*Redirect Recombineering Technology™*" (Gust et al. (2003) Proc. Natl. Acad. Sci. USA 100(4): 1541-6). Se diseñó un conjunto de cebadores para amplificar la casete de *apramicina-attB* usando el cebador3 directo (SEQ ID NO: 10) y el cebador3 inverso (SEQ ID NO: 11). Estos cebadores fueron diseñados para contener secuencias 5' que comparten homología con el grupo de genes de espinactina. El producto de PCR resultante fue integrado en el clon de cósmido 4G15 siguiendo el protocolo de "*Redirect Recombineering*". Un bloqueo del grupo de genes de espinactina en el clon de cósmido 4G15 dio como resultado un doble cruce, en donde la casete de *apramicina-attB* reemplazó la secuencia de ADN nativa del clon de cósmido 4G15. La eliminación resultante fue confirmada mediante secuenciamiento de ADN de la región de unión. La secuencia de nucleótidos de la unión, en la que se había eliminado el fragmento de 7,6 Kb del grupo de genes de espinactina, se proporciona como SEQ ID NO: 12.

El ADN de cósmido 4G15 fue aislado y transformado en una cepa donante de *E. coli*, la S17-1, y se usó en un experimento de conjugación con la cepa de *S. spinosa* NRRL 18538. La conjugación se completó siguiendo un protocolo que había sido modificado para incluir medio ISP4 (Matsushima et al. (1994) Gene 146(1): 39-45). Se seleccionaron transconjugantes inundando las placas de medio de conjugación con apramicina (50 µg/mL) y ácido nalidíxico (25 µg/mL). Los transconjugantes procedentes de las placas de conjugación fueron parcheados en medio agar ISP2 suplementado con CaCl₂ 20 mM y apramicina (50 µg/mL) y ácido nalidíxico (25 µg/mL). Los bloqueos de doble cruce (resistentes a apramicina y sensibles a antibióticos de canamicina) fueron seleccionados en medio agar

ISP2 que contenía apramicina a una concentración de 50 µg/mL.

Se seleccionaron cinco cepas bloqueadas independientes (marcadas como 538-M2, 538-M4, 538-M6, 538-M7 y 538-M10). Estas cepas bloqueadas fueron confirmadas molecularmente para la eliminación de 7,6 Kb del grupo de genes de espinactina mediante amplificación de PCR. Las reacciones de PCR produjeron un fragmento del tamaño esperado, lo que indicó que se había eliminado un fragmento de 7,6 Kb del genoma de la cepa de *S. spinosa* NRRL 18538 (ver la Figura 6). Habiendo confirmado las cepas bloqueadas de *S. spinosa* NRRL 18538 a través de ensayos fenotípicos y genotípicos, las cepas fueron propagadas como cultivos de fermentación en matraces agitados y se cuantificaron los títulos de producción de espinosina A y D.

Ejemplo 6: Producción de espinosina A y D y confirmación de la eliminación de producción de espinactina en las cepas bloqueadas de *S. spinosa* NRRL 18538

Las cepas bloqueadas derivadas de la cepa de *S. spinosa* NRRL 18538, que fueron generadas para contener una eliminación parcial del grupo de genes de espinactina, fueron propagadas como cultivos de fermentación en matraz agitado de 6 días y de 10 días. Las cepas bloqueadas produjeron niveles de espinosina A y D comparables a las cepas de control que contenían un grupo de genes de espinactina intacto. Además, los niveles de espinactina en las cepas bloqueadas de *S. spinosa* NRRL 18538 fueron indetectables o muy bajos. Estos resultados indican que la eliminación del metabolito secundario, la espinactina, no afecta de forma negativa a la producción y biosíntesis de las espinosinas A y D.

Se analizaron extractos de metanol del caldo de fermentación procedente de la cepa de *S. spinosa* natural NRRL 18538 y de las cepas bloqueadas de *S. spinosa* NRRL 18538 para determinar la producción de espinactina usando un método de LC-UV-MS. Se completó el protocolo de extracción con metanol en el que una parte del caldo de fermentación y tres partes de metanol fueron incubadas durante una noche a temperatura ambiente. Las muestras de los extractos de metanol fueron sometidas a cromatografía usando una columna Agilent Eclipse Plus C18 (100 x 3 mm; 1,8 µm) a 40°C eluída a 0,5 mL/min usando un gradiente lineal de A:B 50:50 (0 minutos) a 5:95 (12 minutos) donde A = acetato de amonio 25 mM y B = acetonitrilo-metanol (80:20). La espinactina, que eluyó a aproximadamente 8,8 minutos, fue detectada tanto por absorbancia de UV a 227 nm como por espectrometría de masas de monitorización de ion seleccionado (SIM-MS), monitorizando el ion a m/z 437,2. Se llevó a cabo la cuantificación usando un método de patrón externo con curvas de calibración establecidas a partir de una muestra pura de espinactina. En la cepa de *S. spinosa* natural NRRL 18529 y en las cepas bloqueadas de *S. spinosa* NRRL 18538, se usó el método de UV para la cuantificación de la espinactina cuando estaba presente en concentraciones por encima de aproximadamente 2 mg/L, mientras que el método SIM-MS se usó cuando el compuesto estaba presente en el rango de aproximadamente 0,02 – 2 mg/L.

Los títulos resultantes de espinosina A y D y de espinactina correspondientes al caldo de fermentación de la cepa natural de *S. spinosa* NRRL 18538 y de las cepas bloqueadas de *S. spinosa* NRRL 18538 se resumen en la Tabla 2. Estos datos demuestran que una eliminación génica parcial dentro del grupo de genes de espinactina no tuvo ningún efecto sobre el título de producción de espinosina en las cepas bloqueadas de *S. spinosa* NRRL 18538 en comparación con la cepa natural de *S. spinosa* NRRL 18538.

Tabla 2: Sumario del análisis de HPLC de los extractos de metanol correspondientes a caldo de fermentación de 6 días y de 10 días de la cepa natural de *S. spinosa* NRRL 18538, presentada como “538” y de las cepas bloqueadas de *S. spinosa* NRRL 18538, presentadas como “538-M2”, “538-M4”, “538-M6”, “538-M7” y “538-M10”.

Día	Cepa	Título de compuesto (mg/L)			
		Espinosina A	Espinosina D	Total (Espinosina-A/D)	Espinactina
6	538	232,05	23,70	255,75	0,23
6	538-M2	269,68	31,96	301,64	0,11
6	538-M4	222,17	26,41	248,58	0,00
6	538-M6	265,62	33,57	299,19	0,00
6	538-M7	233,58	29,75	263,43	0,00
6	538-M10	274,41	35,38	309,79	0,13
10	538	213,18	33,87	247,04	5,36
10	538-M2	186,93	27,03	213,95	0,14
10	538-M4	184,25	31,42	215,67	0,00

10	538-M6	258,14	45,06	303,20	0,09
10	538-M7	239,61	43,79	283,40	0,00
10	538-M10	145,07	27,46	172,53	0,08

- 5 Se eliminó un fragmento de aproximadamente 7,6 kb de la secuencia codificadora del grupo de genes de biosíntesis de espinactina que incluye el gen de *tiazolinil imida reductasa*;, neutralizando con ello la producción de espinactina en la cepa de *S. spinosa* NRRL 18538. La posterior fermentación de las cepas bloqueadas dio como resultado la biosíntesis de las espinosinas A y D en niveles comparables a las cepas de control de *S. spinosa*. Por lo tanto, la eliminación del fragmento de 7,6 kb del grupo de genes de biosíntesis de espinactina en la cepa NRRL 18538 de *S. spinosa* dio como resultado la eliminación del metabolito secundario, la espinactina, del cultivo de fermentación sin afectar de forma negativa a la producción y la biosíntesis de las espinosinas A y D.

Listado de secuencias

- 10 <110> Dow AgroSciences, LLC
Palaniappan, Nadaraj
Lewer, Paul
Graupner, Paul R
- 15 <120> Identificación y caracterización del grupo de genes de biosíntesis de espinactina procedente de *Saccharopolyspora spinosa* productora de espinosina
- <130> 2971-P10917US (71553)
- <160> 12
- <170> PatentIn versión 3.4
- 20 <210> 1
<211> 64118
<212> ADN
<213> *Saccharopolyspora Spinosa*
- <400> 1

ES 2 661 447 T3

cgatttgacc	ttcgcggaac	gactcgcccg	cgcgctgcac	acacgcatgt	ggacaacacg	60
agccgacgtg	attcaaaccg	tggacaaccg	tgtccacgcc	acccagggtca	ccgtcgccga	120
agacggcacc	ttggtgccga	ccttcgtcaa	cggcctgggc	accgggcact	ggcacctgct	180
gggctccgaa	ggccaacacc	tatggagatc	cgccccggaa	ctgggtaccg	ccgtcacccg	240
ggacggctcg	acacccccgc	gttacaccga	cggaacgcac	cccgccccgg	taatcagatg	300
ggcgggaagc	aacagcggcc	actctggaca	acaggatgcc	gaagcggaaa	cggcgaagca	360
gactcggcgg	cccgggatag	agaaggatga	gttgaacgag	tgggtggagc	gggctaagaa	420
cgaggccaag	gccgaaacgg	cagcggcagg	actgaaatgg	ttgaaaggac	agggggccgg	480
ggttaataac	gaggtttggg	ggaaggcatg	gcttgctgcg	gatgggaagg	agaagcaggt	540
ccggcggcct	ggggtagata	aggagcgggt	gaaggagtgg	gttgagcggg	ctaaaaacga	600
gggagggggt	acgtcggcag	cggcaggact	gaaatgggtg	cgtgaaaagg	ggatcgggggt	660
tgggaaaaac	gtttgggggg	acgcgtggct	tgctgcgact	gggaaggaga	agcaggttcg	720
gcggtctggg	gtagataagg	agcggttgaa	tgagtgggtt	gtgcgggcta	aaaacgaggg	780
aggggcccag	acggctgcgg	caggactgca	atgggtgcac	ggacagggga	tcgaggttaa	840
caaaaacgtt	tgggggggatg	catggcttgc	tgcggtggg	aaggagaagc	gggttcggcg	900
gcctcagcta	gataaggagc	ggttgaatga	gtgggttgtg	cgggctaaac	acgagggagg	960
ggccgagacg	gctgcggcag	gactgcaatg	ggtgcgtgga	cagcggttcg	cggttaacac	1020
cgaggtcttg	ctggatgcgt	ggcgtgatgc	ggccgggaag	gagaagcagg	ttcagcggcc	1080
tcagctagac	aaggagcgggt	tgaatgagtg	ggttgggcgg	gctaaaaacg	agggaggggc	1140
caagtcggct	gcggcaggac	tgcaatgggt	gcacggacag	gggttcgggg	ttgacaaaaa	1200
cgtttgggcg	gatgcgtggc	gtgctgcggc	tgggaaggag	aagcaggttc	ggcggcctgg	1260

ES 2 661 447 T3

ggtagataag gagcggttgt ttgagtgggt gttgcgggct caacaggagg gaggggccga 1320
 gacggttggtg gcagggctgg aatgggtaaa tcggcagggg gtcggggttg agaaaaaggt 1380
 ttggtcggag gcgtggcgtg gtcgtggggt tttggattac tggcctgctg attttgacgg 1440
 tgacgaggat ccattcgtgt gggaaagggg tgttggtgag ggcgatgtcg ccggcgttct 1500
 ggatgtgtcg ccggtggatt gggagaggtt taccgagttg gaggtcggc cggaggccga 1560
 ggtggcgagg ttgcagtcgg ctgctgcgga gctggtgagt ccgctcttgt tcgctccgca 1620
 gtttgtgggt gtgaagccgc cggagcgggt ggcgtttgat ttggtggtgg atcaggttgc 1680
 gatcacgctg catcgagcct cgcgagatgg tctgcccgaa ctggaggccg gtgaggccgc 1740
 gcggacgcgt gccgaagagc tggcggcaag ggtgcgcgag gtgggtttgg aaggcgcgcc 1800
 ggatatgcgg cgcactgcct caccgccagc acgtggtgtt gaggatgtgt cgtcctctgt 1860
 tgggggagtc gggcagcgcg aaccgggtggc ggtgggggat gcggcctcgg aacgggggtc 1920
 ggcggatgctg tatccgggtg agtcgcagcc tgggccgtcg acggtccagg ccaccgccca 1980
 ccaggggtcg gctgcgggtg atgtggccgc ctccgggcag caggatgctg atgcgggaac 2040
 ggcgaaacgg gttcagagcc ctgcgataga taaggagcgg ttgaaggagt gggttgtgct 2100
 ggctcaacag gtgggagggg cggagacggt tgaggcagga ctggagtggg tgcacgcgca 2160
 ggggtttggg gttgataccg gggcttgggg ggaggcgtgg cgtggtcatg gggttctgga 2220
 tcactggtct cctgattttg aggatgaggg ggatccattc tcgtgggagg ggtttgacgg 2280
 tgttggtgag cgggatgtcg ccggcgatct ggatctgttg ccggtggatt gggagaggtt 2340
 tgccgagtcg gaggtttggt cggaggtcga ggtggcgagg ttgcagttgg ctgctgcgga 2400
 gctggtgagt ccgctcttgt tcgctccgca gtttgtgggt gtgaagccgc cggagcgggt 2460
 ggcgtttgat ttggtggtgg atcaggttgc agtcacgctg catcgagcct cgcgagatgg 2520
 tctgcccgaa ccagaggccg ttgcggcggc gcggacgcat gctgaagagc tggcggcaag 2580
 ggtgcgcgag gcgggtttgg aaggcgcgcc ggatatgcgg cgcactgcct caccgccagc 2640
 gcgtggtgtt gaggatgctt cgtcgtctgt tgggggtgtc gggcagcgcg aaccggtggc 2700
 ggtgggggat gcggcctcgg aacgggggtc ggcggaggtg tatccgggtg agtcgcaggc 2760
 tgggccgtcg acggtccaga tcaccgccca ccaggggtcg gctgtgggtg atgtggccgc 2820
 cgctgggcag caggatgctg atgcgggaaa ggcgaaacgg cggcctgcga tagataagga 2880
 gcggttgaat gagtgggttg tgcgggctca agaggaggga ggggccagga cggctgcggc 2940
 aggactgaga tggatgcgtg gacaggggct tgggcttgat agcaaggttt ggggggatgc 3000
 gtggcgtgct gcgactggga aggagaagca ggttcagcgg cctggggtag atagggaccg 3060
 gttgtatgag tgggttgtgc gggctcaaga ggagggagg gccaggacgg ctgcggcagg 3120
 actgcaatgg gcgcgtggaa aggggttcgg ggttaggaaa aacgtttggt cggaggcgtg 3180
 gcgtgctgcg gctgggaagg agaagcaggt tcggcggcct ggggtagata aggagcgggt 3240

ES 2 661 447 T3

gtatgagtgg gttgtgcggg ctcaagagga gggagggggc gagacggttg cggcaggact 3300
ggattgggtg catcggcagg ggttcggggt taacaaaaat gtttgggggg aggcgtggcg 3360
tgctgcggct gggaaggaga agcaggttca gcggcctggg gtagataagg agcggttgaa 3420
tgagtgggtt gtgcgggctc aagtggaggg agggggccgag acgattgagg caggactgga 3480
ttgggtgaat cggcaggggt tcgggggttaa caaaaatggt tggggggagg cgtggcgtgg 3540
tcgtgggggt ctggatcgct ggtctgctga tcttgaagat gagggggatc cattctcgtg 3600
ggaggggggc gacggtggtg gtgagcggga tgttgccggc gatctggatt tgttgccgac 3660
ggattgggag aggtttgccg agttagcaca agccgggcac gatctctctc ccgtcgcaca 3720
cgatctgcac gacacgctgg accgtacgca caaggacctc cccagcgacg cccgttccac 3780
cgacgacgtc ggactgaccg caccgccgcc gtttccgcac gccgaaggac cggcacgggc 3840
ggagatctgg atgcagccgg gcagcgacgg ggtgaaaccg gtaccagaca cgatccagga 3900
ccgacgcccc aaacgagacg gcgcacgacg gcgcctggcc tggcccggta gtaggcgggc 3960
cgacgacacc gccccacca ccgcatcggg ttcctccac gcgatctcac gggcttcccg 4020
ggaagcccggt gagatcccca gtggggcggt tcgatcagg ccacaggaaa acaccgggca 4080
ggtcacccgg agggacggtt tcgctggttt ccttggtggg tacagcgacg accacgggat 4140
gggttcggtc gggctgtctg acgaaagcac tttcggtgac ccgtcaccgg aagctgcggc 4200
gtggccgggt gcggggcggt tgggtgggtc tgggttccgg tttgtgcggg gggatgaatca 4260
ggccaattac ttctttggtg atgcgcggtt tcgggtgaat tgcctggagg cgtgtgtggc 4320
gtttcataat tcggtgaagt tcggtcggca gttcgtggcg ggtccggccg gtgctgatcg 4380
ggatccggct cggttggagg tggcgttcgg ccggcaggcc cggcaggtag aggcgtgtggc 4440
cggggttag cggatatgtc gtggcgggcc ggtggggggc actgtgccgg tgctgtacca 4500
gcgacccgac ggcagcgcgc atgtgatcaa tgcggtgctg accggcagca aagaccctgc 4560
tgggcgcgggt gtggtcgtgt tgtgggatcc gcagcggggc gaggaagccg aggtagccga 4620
tgtttctgcg gtgacgggga tgtgggtgat cccggtgccg gaggcggaag cagcggggat 4680
ggtggtggtt ccgtccagtg gttccgggtt gcgtagccag ggctggggat ccggggttgc 4740
caacgagggt gcgggtccga agcgtcagcc cgggtcgtcc gggccggagg ccgggaaacc 4800
cggcgcgggc gcgaaaagga cacagcgga gaccggtgag agttcgccgg gtgccaagcg 4860
gcggaagggg gcagcagccc cggctgggat tgaccccggg agtccgggca atggcgaggc 4920
tggggtgctc accccgaccg atcaggaac acaacaggac gagcgggcca cggagcgga 4980
agagaaaaag aaggcaaggg acgcgaagta tcgccaggcg agaagggccg aggccgatcg 5040
tgttgtggtg ttggaggagt tggagggggc ggggcagttg actgaggagc aggcgacgga 5100
gctggcggag ctccagccga aggtgacgga gcggaagcag caaaagaagg aagagaacgc 5160

ES 2 661 447 T3

gaagcagtac aaggcgagaa aggctgccgc tgcgcgggtt gtggagttgg aggagttgga 5220
 ggggcggggg gagttgactg aggagcaggc gacggagctg gcggcgctcc agccgaaggt 5280
 ggcgcaacag acgcagcaaa cgaaggaaaa taacgcgaag cagtaccagg cgagaaaggc 5340
 tgccgccgat cgtgttgccg tgttgagga gttggcgggg cggggggagt tgactgagga 5400
 gcaggcgacg gagctggcgg cgctcaagcc gaaggtagcg cagcgaagc agcaaaacga 5460
 aggaaaagac cgcgaaactat agccaggggg taaaggctgc cgatcgtatt gcggtgttgg 5520
 aggagttggc ggggcggggg ccgttgactg gggagcagga ggccgagctg gcggaactcc 5580
 agcgcaaggc ggcgcagcgg aagcagcaaa agaaggaaaa gagcgcgagg cggtagcagg 5640
 cgagaaaggc tgccgccgat cgggttgttg tgttgagga gttggagggg cggggggagt 5700
 tgactgagga gcaggcgacg gagctggcgg cgcttcagcc gaaggcgagc cagacgcaga 5760
 aaacgaagaa agataacgcg aagtatcgcc aggcgggcaa ggctgctgcc gcgcgggttg 5820
 cgggttgtga ggagttggcg gggagggggc cgttgactga ggagcagcgg acggagctgg 5880
 cggcgctcca gccgaaggtg gcgcaacaga agcagcaaaa cagggcaaag taacgcgaag 5940
 tatgctcagc gggtaaacgc tgctgccgat cgtgttgttg tgttgagga gttggcgggg 6000
 cggggggagt tgactgagga gcaggcgcg gagctggcgg ggctccagcc gaaggcgagc 6060
 cggaagcagc aaaggaagc aaggagctg aagtataacc aggcgagaag ggctgctgcc 6120
 gatcgtgttg cgagttgga ggagttggcg gggcacgggc cgctgactga ggagcagcgg 6180
 acggagctgg cgcgcctccg gccgaaggca gcggggcggg ggccgagcgg gaagaagaag 6240
 gaccgggagg tgatggagac cggggtgggt gaggcccgcg tggcgggcg ggatgagcgg 6300
 attgcggggc cgaggggggt atcgagtggt actgggactg atcagggaag tcgggacgca 6360
 tggtcggctg acttcgatct cgtgctggc cccgagcagg cgggtggcggg ttcagcttgg 6420
 tcgggggggt tgggtgcggg agatgccggg gtgatgctgg gcgagggcgc tgtcgaggac 6480
 gttctcgca ccgagttgac cgcgttcctg ggacaggatg ccggagcggg tgatttcgcc 6540
 ggattcctcc ccgactacca cgactgggac ggctccgttg ggttgttcgg ggccgaaggg 6600
 cctgacaggg gacctttcgg tgaccattt ccggaaggtg cgggtgtggc gaatgaggtg 6660
 gctgcggatc cgggtgggtg gatggggcg tctggggtgt gggcggattc cgaccgggat 6720
 gcggacggcg cgggtgggtc tgggttcggg tttgtgcggg gggatgaatca ggccagctac 6780
 ttcttttgtg atgcgcgttt tcgggtgaat tgccctggagg cgtgtgtggc gttccataat 6840
 tcggtgaagt tcggtcgca gttcgtggcg ggcgcggctg gtgatcggga tccggtcgg 6900
 ttggaggcgg cgttcggcg gcaggcccg cgggtggagg gtgtggctgg ggttgagcgg 6960
 tatgtgcgtg gcgggccggg ggggtggct gtgccggtgc tgtaccagcg acccgacggg 7020
 agcgcgcacg tgatcaatgc ggtgcgtacc gagagcaaa accctgctgg gcgcgatgtg 7080
 gtcgtgttgt gggatccgca gcggggtgag gaagccgaga gagccgatgt ttcagcgggtg 7140

ES 2 661 447 T3

accgggatgt	gggtgatccc	ggtgccggaa	gcggagggcg	aaccggacca	ggggatggtg	7200
gtgttgccgt	ccagtggttc	cgggttgccg	agccagggct	ggggctccgg	gctgcccacc	7260
gaggtggcgg	gtccgaagcg	tcagcccggg	gcgtccgggc	cggcggccgg	gaaaaccggc	7320
gcgggcgcga	agaggacacg	gggcgaggct	gatgagagga	catgggcgga	gaccgggtgag	7380
agttccgggg	gggccaagcg	gcggaagggt	gcagcagccg	cagctgggat	tgaccccggc	7440
attccgggca	atggtgagag	tgaggtgctc	accccggccg	atcaggcagc	acaagaggcg	7500
gcgcagcaga	gacagaaaaa	gaaggaagag	aacgcgaagt	ttctccaggc	gagaaaggcc	7560
aaagtgcctc	ggtttgccga	gttggagaag	ctgaaggggc	aggggcagct	gactgaggcg	7620
caggagacgg	aactggcggc	actccagccg	aagggtggcg	attggaaaca	gctagagaag	7680
gcaagaaacg	cgaaggatca	ccaggcgaga	aaggctgctg	ccgcccgtgt	tcgaggagttg	7740
gaggagttgg	cagagcgggg	gccgttgact	gaggcgagcg	cggcggagct	ggcgaagctc	7800
cagccgaagg	tggcgcagcg	gaagcagcaa	aagaagaaa	atagcgcgaa	gcggtcccag	7860
gcgggaaagg	ctgctaccgc	ccgtgttgcg	gagttggagg	agttggcggg	gcggggggag	7920
ttggtgaggg	cgcaggcggc	ggagctggcg	gagctccagc	cgaagggtgc	gcagcggaag	7980
cagcaaaaga	aggaaaatag	cgcgaagcgg	tccagggcgg	gaaaggctgc	taccgcccgt	8040
gttcgaggag	tggaaggagtt	ggcggggcgg	ggggagttgg	ctgaggcgca	ggcggcgag	8100
ctggcggagc	tccagccgaa	ggtggcgag	cgggaagcaga	aacagaagga	aaagaacgcg	8160
aagcgggtccc	aggcgggaaa	ggctgctgcc	gcccggttg	cggagttgaa	ggatcaggag	8220
gagctgacta	aggaggaggc	ggcggagctg	gcgaagctcg	aggcggagggt	ggcgaagcag	8280
agggaaagggt	ggaggaaagc	tagcgcgaag	catcgcccgg	cgaaaaaggc	cgccgctgcc	8340
gccgcgattg	cggagttgga	ggcgttgga	gagcagggga	cgctgacttc	ggagcagcag	8400
gccgagctca	atttcggcg	ggaaataaaa	atgaacgaaa	ctgagggtaca	tcgcctggta	8460
aaccagggtat	cgctagggaa	caatgccgct	aaggaaagtca	aggcgctgga	ggcgttgccg	8520
cggagtggtg	agattgaggc	caagctgcct	acgttgcgga	gcaagggtgc	taggttgccc	8580
gcgtttgcgg	agccgttgaa	ggaaacgaaa	caaaggcttg	cggaaagcag	gaagaagctc	8640
aagctgatgc	aggaggctgg	atcggggcgg	attgcgggtt	tggaggcaag	tgggcagcag	8700
gatgagcaga	gcgcaatggg	ttcgagcggg	gtatcgaggt	ggactggggc	tgatcgggac	8760
gagcggggcc	agggatggc	cgacttcgat	ctcgggtcgt	ggcttgatcg	ggagctggcg	8820
gatttgatg	tgctcgtggg	tcgggtccg	gcggagcagc	gtggggcgga	acgggacgcg	8880
tggtcgggtg	ggggcggtga	tctcgtgctg	tcgggtgatg	aggcgatcga	ggattcagct	8940
ggcgacgggc	aggctcgggc	atcgccgtcc	actttggacg	aaaatcggca	gacggatttg	9000
gaggcggtgc	gggagtcgga	tcgggtatg	cagacggagt	ggggggagcc	gcgagagggg	9060

ES 2 661 447 T3

ttggacgagg ctgatcaggg gcaggtggag ggtgcgatcc gcgatgcggc gcaggattcg	9120
cgggttgccg agctgcggga gtcgctgaag caggctgagt ctgacctggc ttcgttgag	9180
gcgttgccgg gttgggacga gaacctgcgg gcggagttcg cgacgttgcc gatgttcgat	9240
cagggtttgc ggacggactg ggagcaggcg cagcggcgta tcgaggcgtt gcgggcggaa	9300
ttggccgaga tggtagatga cgcgggcgaa ccggccgggc aacacatggg tagtctcgg	9360
gaggggcggg ctgcggcatt cgtgccgact gagattgcga cgttggtga ggtgacgggt	9420
cggatggagc ggcatattcg gacgatgccg ctggaattca atccgaatcg tgtgcgtgag	9480
aacgtcgaaa ctgcgaagca gtgggtgcag caggggcagt ggactccagc gcagttgcgg	9540
acggttgagt cgaggttgga ggcgatacat gcggcgaga gggcgttgct cgatgctcgc	9600
ctggcaatgg atgcgcaggc gataggcggg gctgagttgg gggctgggta tcggttcctg	9660
gagggggcga accgggccaa ttatctctct ggtgatgtgc gtttccaggt gaattgcctg	9720
gaggcgtttg tggcgtttca taattctctg aagttcaacc gggagtttgt ggcggggccg	9780
gccggtactg atcgggatcc ggtcggttg gaggtggcgt tcggccggtg ggcctggcgg	9840
gtgggtggtg tggctgaggt tgagcagtac gtggggagcg ggcgggtggg cgttgctgtg	9900
ccggtgattt accagcgagc cgacggtagc gcgcatgtga tcgctgcggg gcacgcccga	9960
gaccgctatg ggcatcaggt ggtggatctg ctggatccgc ataaggggga ggtcgctgag	10020
aaagctgatg tcttggtgc gaccgggatg tggatgattc cggtgccggg gccggtgccg	10080
gtgccggtgc cggtgccggg gccggtgccg gaccgggaga tgggtggtctt gccgccagc	10140
gggtccgggt tgccgagcca cggctgggga tccggactgc ccaccgaggt gacgggtccg	10200
aagcgtcagc tcggtgcgtg gccggtgacc gggaaaaccg gcccgaaagg cacaagagg	10260
acgcaggcgg aaggccgatg agagtgtcgg ggcggcggct ggcgagagtt ccggggccca	10320
ggtggcagca gcagcggctg ggattgatcc agggagtccg ggcaacggcg agagtgaggt	10380
gctcactccg accgatcagc cagcacaaca ggacgagcgg gcggcgagc ggagacagaa	10440
aaagaacgaa ttgaacgca cgtatcgtg ggcgaaaagg gccgaagccg atcgggttgc	10500
ggtgttgga gaagctcaag ctgatgcagg aggttggtgc gggtcggatt gcgggcgtgg	10560
tggcaagtgg gcagcaggat gcaatggtgc taggtgaggt gggggagtgg actggggctg	10620
ggcagggcga acaggacgcg gggccggctg acgtcgatct cgatgggtgg catgatcgga	10680
cgttgccgga gttggatgag ggcagggcgg atgcggtgga gggcacggcg gatgctggcg	10740
tgatgcaggg cgaggatgct gcttacgagg agtttgcgc gaccgcgttg accgcgttcc	10800
tgccagagga tgctcgggag ggcgcggcgg atgctgtgga gtccggcaat gaggcaatga	10860
gggctgggta tcggcacaca atcaggctgg aaaaccacc acagcgtgaa taatgcagg	10920
atctccaagg atgcccgcca gcttcggcag aggcggatca tcgcgagtgt gatcatggtt	10980
tcggagttgg cggtcaggcg ttcgtcgtct cgtgcgaggc ggcgatgtcg gatcaggcac	11040

ES 2 661 447 T3

ccaaaagggtc gttcagaccac ccacctgcgc ggtaggccctt ggaaccctt gacatcgtcg 11100
 ctgcgtttga cgatttccag taccagtcgc aggttttctt tcgcccaggc cggcaggccg 11160
 ttgatgatgg agttggggtg tccggcatcg gtctggacca gcgcgatcgg gggaaccggg 11220
 cggccagggtg ttgcaggatc tccgccttc cgacgtggtc ctgtactgac gcggagggtga 11280
 ccatgacgggt cagtagcagg ccgaggcggt cggccgaggg tgtccacggg gaggtggcgt 11340
 ttgcggcctg tcgtcttctt atcgaaccgc tgtgcttcgc cgcgctcgct gctttcgatg 11400
 gactgggcat ccagcaccgc cggcgtcagg tccgcctggt gcccgccctg gatacggacc 11460
 tgatcgcgca gcgcgtcgtg aacacggtcg atggtgccct cgtgtgttcg ctacgtgccc 11520
 acctggttac ccaacagtct cttggcttgc ctctaaaaag ttaccctggt caaactggcc 11580
 gctgtgtgat cacggaacgc cttgggtgat catttttctt gccacaact caagcggcgc 11640
 cggcatgacc ccgtcccggg gcaggggcag tggctggcca gtgtcttcg agagcacacg 11700
 gcctactacg ccgtaccagg caactccgat gcgtatcgac ctcccgccat cagggtgacca 11760
 ggctctggca caagcgctc cggcgccgaa gccggacatc cctgggtggg gtgtgaaccg 11820
 cattgttgac cgcgaatgat cggcacgaac accaggatga tcgcctgcat cagcacagtg 11880
 atgaggttct gcggttaggt cagtatgctc gccgcagggt cgtgatcga gtggtaccag 11940
 tggccagccc gtgtcactcc agcgcgcgaa ggttcggcac gacgtccacg acggagtccg 12000
 gtaccagagt gccgcagccg ttgaaccgtc gcggcctgga ttacgtgtgg ctacaaaaac 12060
 gctgatcggg gcccatacca tcggtgctgc agggcggaag gagtggtaga tggcgagctt 12120
 attgtctcgt cggacgcaaa ccgctggcga tgatgcgctg atccggctgc tgttcgagga 12180
 gcacggctgc gccctgctcg ctactcgac gcgactgacc ggcgatcag cggcgccgga 12240
 ggacgtcgtg caggagacgc tggtagagc gtggcgaaat cccaacgcc ttgtgaaccg 12300
 gaagggttcc gttcgcgggt ggttgcgtgc cgtggcgcgc aacatcgtca tcgaccgcag 12360
 ccgcgcgaag gccgctcgtc caccggagggt cggcgccacg ccgacagcgc cgcgggcagt 12420
 gcgtgaccac gccgactccg tggtagactc gatggtcgtc accgaggcgc tggatacgt 12480
 ttcggaagac caccgcgacg tgctggtgca gatctacttc cgaggactca gcgtggctga 12540
 ggcagccgct gaactggcaa taccgcccgg aaccgtgaaa tcacgatcct actacgcctt 12600
 gcgggcaactt cgtgaagcgt tgggcggcgc agcgttggtc gcatggaaag aggtggctcg 12660
 atgagcagcc aggatcacaa ggtcgacctg ggcgcctacg tgttgggctg actcgacgaa 12720
 ccggaacaca gcgcgctcga cgagcacctg acgtcctcgc gacactgccg tggggaactg 12780
 gaggacttgg agcagatgcg ggccgtgctg gacgaggtcc cgcgggagtt gttcctggac 12840
 ggacctccgg acgacgcaga cctgctgttg cagcgacact tgcggcagggt tcgttcggag 12900
 aaatccaggg cgtatggcat caggcacgcc gtggtaggcg cggcgccggc tctcgtcgcg 12960

ES 2 661 447 T3

gccgtcgccc tcggcggcgg cgtgctgato ggccgagcgg gtgacagcgc accgccgctg 13020
 gtcgtcgcgc agccgccgat cgcgtccatg ccgcccgcgg gaacgaaggc cgccacggcc 13080
 gtcgactcgg ccagtggggc gaggtgact gtgacgtca cgcctgcagc tggctgggtg 13140
 cggttgaacg ccgcggtcaa cgggataccg gcgggcgaac ggtgccgctt ggtcgtagt 13200
 ggcaaggacg gtagccggga gatcgctggt agctggctgg tgtccgagaa gggcgcgaac 13260
 gaaggcagcg cgctggatgg ttccggcgtg atggcgccgg acgatgtggc ggccgtcgaa 13320
 gtggagaact tcgagggcac cagattcgtc tccacggcgg tctgaccagg cgacatattc 13380
 cgacgtcaag ccccgctccg gaacatccgg ggccgggctt gctcaggccg cggcccggga 13440
 gtgcgatgat ggctcggtgt ccacaagcca aagacagtct gggcgccggg cgtggcggct 13500
 ggccgcagtt ctgctgccgg tcgggctcgt cgcggcgttg ggtgtcgggt acgtcgacgg 13560
 gcacgggtgt catggcgctc cggggcgggt gaccggcgat ttcgctcgaca tcgcagatgt 13620
 tcccccgatg ggtccgcaaa accggttcgg ccgggacgcc tcgaccggcg tgatgacggt 13680
 tgactgcggt cgtaacgaaa acgcccacgg caacgcggac aacgtcatcg catcgctgc 13740
 cgtgtccggc ggagcgcac acatgcacga ttacgtgggc aacctggcaa cggatgcgtt 13800
 cgccaccgac cagagcctcg ctgcagccga aaccacttgt gtgaacggcg acaggtccac 13860
 ctactactgg ccagtgcctc gcctgctgaa cggcgacgag ccgggttcgg ccgatatgcc 13920
 gcacaaccgc ggcacgattc ttcaaccggc ctcagtactg ataagtttcc ggggcaatcc 13980
 ggccagcaag gtgatccga tgcgccggtt cctgcgggcc gtcacgcggc acgcgaaggc 14040
 gctcaccaac ggcttggaac acgccgcccg agtgcaactg acctgttcgg gaatgctcga 14100
 ccggcagacg aaccgctatc cgaggtgccc ggacgggcat cgagtgttcc gggctctcga 14160
 ctttccgagt tgttgggatg gcaggaggaa cgacagcccg aaccatcgcg atcacctggt 14220
 gttcccgcg ggaaacgggg tatgcccgc caactcgttc gccgtcccgc agttgcgcat 14280
 cgaagtcgac tacgacgtgc ccagcgacgc ggagtacgcc atcgacacct tccccgatca 14340
 gcagcgcagc ccgatcacgg atcacagcga tttcataaac gtgatgaccg atgagcagat 14400
 ggcacacgtc gtggcgcgtc tgaacaacgg gcgcacggc tgaggaagaa acctccgaaa 14460
 tgaaaatggt cgtgagtgcg ttgtccgggc tccgcccggc aacgcactca cgaccattgg 14520
 tcagcggtag gtcagccgaa gaacaccgac tccacggcat tgtcgaatcc gacctcgccg 14580
 agatcggcga catcacctgt gatgaccaac gacttgccgg tgcagttctg cccggaccac 14640
 agcttcaagg acccggaacg cgaatcgac gacgtcacgt cgcgccgcgc gacgttcttg 14700
 cagcccggtc cggccacgcc ctgcgaagca ccgtcgtcgc tgaaattctt gtcgtcgaag 14760
 agaaccgcgc tgtcggcgga ttcggtcggc tgcgtgctgc cgtcctgcgc atcgccgttg 14820
 ccgtcaccgg tttcgggtccc ggctttctgg ccgttcggag tgatcccgaa ccacgttccg 14880
 cccacacctt ggccttggc gtcgccggc ttggtgtcct tggcgaagcg gtagaccggc 14940

cagccaccga tgggtcacctg gaggtgccc tegtcccgtt tgacgggtgcc caccttggac 15000
 ttccgaacgc cggccaagaa gatcctgccg cccggagcca cgggtcaacgg cggccaggtc 15060
 ttccgcgaat ccccggttga gttcgacacc gagggcttgg ccgtgtcctt gtcgaaccgg 15120
 tacaacgtga gtccggcacc gttgaccact accgggtcca gagcacctgc ctgggaagcc 15180
 ttcaactgca cccactttcg cgcaacattc tggtcgcggg tgttggctgt gccgtcggac 15240
 ccggttgccc agtccccggg cttcgccgtc gcattgtgcg ccttggcggg ccccgccagt 15300
 aaggacaggc cggcgcttcc ggtctcgggg acctgcgcac tcggcgcggc aggcgcggga 15360
 cccgcgcga aggtcggcaa cccgcacgcg ctcaacgaca ccgtccggc gacgacgagg 15420
 acggcaacgt acttggctcg ctttgacagc atgagttctc ccctcaaata ccggggtcca 15480
 gccaatctgc agccacgtac ggcggaacaa cgtgccccga tccgcaccgg ttcaagagat 15540
 cgctgaattt ttccgacgcc cgccgttgac tcgatggatg gcgtatggtg atcgcgtcat 15600
 gtctgtgagt gacatgtgga cagagcgagg atggacgcg gggcgctggg cagcggcgat 15660
 caaccgcac tttgcttcgg tgcgtcgggc gggcgctcgc gcattaggtg atcggcttgg 15720
 ccgcacgatc accaaggacg atctttggtc gacgctcgt tcagcaagta ctcgatctcc 15780
 gtgagccagc acttcaacac gtccacgctg gcgacgatcg gcggtctcct acggaccggc 15840
 gtggcgctgg cggctttgcc ggccttgagg ttgcagggtt tcggtttga ggatctcgtc 15900
 gcggtacccg tggattcgcc gggagcgacc aggggtcatcg gcgtggtgta ccaactgctc 15960
 caggacctct cgcctgtcgc ggctcgtttt ctggacttca tgcgtgcgac cgagatcgat 16020
 ccccgcgag gagtctcccg cggcctcgac ggcgacaccg gcgaaccaga cgcgatcaac 16080
 tggcaccgta catgatcagc cagcgcgccg atgccggctt gttgcgcgcc gtgtgtccat 16140
 ttctgtcgga ctttccccgg tatgaaggcg ttccctggga cgattccgct tccgcccga 16200
 atctgtccaa ttgcgactct ccagcgaggc ttagtctctg cggttcctat tcctggtgcg 16260
 ccaccagacc acgggagccc tcttgagccc tctacttctt ggcccgagtc accgaccca 16320
 gctgagccgg gggcctgctt agaggctccg gcaataggcc gggaaagtgt attagaaccg 16380
 caggtagaag tccgtgagcc cgttggtgct tgcagaggcc gctgtagaac ccgttctgcc 16440
 tagaggtttc ggtgcgcatc ccaacgggcg cggctctcgt ccgcctcccc aaccgatgtg 16500
 aagcgaacgg cctgttcgct tcggtaggtg gcacgaacgg tccgttcgct tcgttcccct 16560
 tgggtgtggg gccggggttg ccggtgtggg gcgaacggcc tgttcgcttc ggtaggtggc 16620
 acgaacggct cgttcgcttc aaacaacacg ccaaacgctg ctaccgcacc cagccgcgct 16680
 cacggtttct gggcggggcg gcgggcgggg ggcgggttc ttaccagggg actcgggtga 16740
 gcttttccgg gttgacgatg tcgcagatct cgtctcgtc gcgttcgctg accagcgcg 16800
 attagcgggc gccctaccag ggcgaccagt gctgcggcca cgtcgggaac cttgaacgac 16860

ES 2 661 447 T3

gccgcgggta cgcgcgcggt gatgtccgct gccgggctgg ccagagcccg tgtggtgcc 16920
 cggcgctcgag ggcggcgtag cggtcgccat cgacgaattt cgttgccaga ccgaaagtgt 16980
 tggcgtaaaa ctgcacggca gcgtcgacgt ccgccacccg gtgttgctgtg ttcccgatth 17040
 tgattgggtc gatcttcgcy ggggcactca tcgggcgacc cgtagccggt tccttctgcy 17100
 ccgtcgaggt gtgcagctcy gcaacctccg cagcgaggat ccggggcagc agtggggttt 17160
 cgagcccgcy gctgctccag acttcggcac ctccgctcca gacttcggca cctccgcytc 17220
 cgttctctcc ggtgaataga ttcttgcyat agttgtggtg ccacctcgcy ccgtacggcy 17280
 cgaacgtggc ctccgcctcc cgtgcacccg tcagcgcggt ggttcttcac gaacctgttc 17340
 ggcccctgca ccagcagcy gcgagccgtg accagcttg tcgcctgctt gtaccgggtg 17400
 agcgcttcgc aaaccaccgc gaaggcatcc acgtcccaag gcgtgagcac tccgaccgcc 17460
 atcagattgg gcgccaaccg gtcccacgcc gcacgggcy cctcgagcy ctccttcaac 17520
 ggggtgacct cgccttcggg cgggatcgcc tgggtgatgc ggtcggtgcy gtcgcygcy 17580
 agcaccttca gcgcgcgatg gctttccgag cccacagcy tcggggagtc aagccaaagc 17640
 gcgcgaagct ggtcgagccg ctgccggacc gccgggtggt cctaaaagtg caggagcagc 17700
 acgggaccgc gctccagccc gccgtctcy taggagaca ggtgggcgga cacctcgtyg 17760
 tggcggggaa gtgcagcggt ttccacgacc ggcggcgtyc cacacggcy tcgccccttt 17820
 gtgcagtggt cgcggaagct cgtgcagcaa aggtttccg gggttgccc ggtttgttc 17880
 cgcgaatatc acagttcaat gcgctggagc ttgccaacat atcgaaagga ggcgcctaaa 17940
 tgcctctccc gcagtaccg cctcaccag gccgcgtyc tcgcccctcy cgtgggcgcy 18000
 atgaccgcy gcggtgagcy cgcgcgacct gaggcgccg caccgttcaa cgcgtcgtyc 18060
 tgcgcgagc tccgcaagag ccaccggaa ctgctcggtg ccgagctgaa ggtcggcagc 18120
 accgttggtc agggcaactt cgaattcatc acccccagc agccggtatg ggtgaagggc 18180
 ctgcaccccy acctggtgag ccttgtyggc gactgcctcy gcttccgtyc cagtytcgcy 18240
 cagatggagt tcaccggcct ggtgcccgc ctccagtyc gacgcytgga cctcgtycgy 18300
 tcgaacctct acgtgaccga ggaacgcygc cgggaagtyc acttcgtyc ctacctgaag 18360
 acgatgagty ccgtcgtyc gcagaagggg aacccgagga agctcacgag cgcgcagcy 18420
 ctgtgcygga ccacggccgc gcaggtgcy ggaaccgtyc agagcgagat catcgagaag 18480
 cagtycgagt catgcygcy gtcgggccc gccgggatca caccgtygt cttctccgy 18540
 ctgggaccag ccgtgcccgc gaccaccacc ggccgcygc acttcttcat gaacgagcy 18600
 ggcatcgca agtcagccgt tgagaagtcc ccgaacacc tcgaccaggy tttcttgtyc 18660
 gacaccaacc tcgtcatgg catcgccag aacaagaaca agagcgagct ccgcygcycc 18720
 gttcacgag tygtgtygcy tgcgaggyg gacggtgcy tgcgcygct ccagcggaag 18780
 tggggcttcy gcgagaatca gctcatgcy cccagatcy tggcctgag gaggttccc 18840

ES 2 661 447 T3

tgccgtcctt ctggcactac ttctgtctgc cgtcgatgtg gcagggcctg ttcaccgggc 18900
 tgaagatcgc cgccgtcgcg ttctgtcgcg gtctggtgct cgggctcgcg ctggccgtcg 18960
 cccgcgacgc cgcgctgggc gtctgtcgcc tgatgagtgc gggctacgtg ttgctgttcc 19020
 gcggtacgcc aatcctgtctg cagctggtct tctgttcaa ctctctccc gccgtcggga 19080
 tcgggctgcc gccgtgcag accgcgatgc tcggcttcgc gctcaacgag gccgggtact 19140
 gcaccgagat ctccgtggc agcctccggt cggtgagccg gacgcagacg acggccgcgg 19200
 cgtcgtcgcg gatgggccg tggctcagc tctcgcgcat cgtcctgcg caggcactca 19260
 agggccgcat cccggcgctc ggtaacgagg cgatctcgct cgtgaaactt acctccatcg 19320
 cgtcagcgat ctcggtcacc gagctgacat tccgcagcca gcagatcgtc gccaccaact 19380
 tcgaggtcgt gccgtatac ggcgtggcag ccgtgatgta cctgctcgtc atcacggcga 19440
 tgaccgccgc gcagtcgcg ctggaacgtc gattcggcgc tccggccgac cgtccgtcgg 19500
 cccgggcgcg cctcttcgcg ttccgcagcc ggacgcgcct tccgcgcgc gccgcgtcag 19560
 cgggtccggt tgccggggtc gaccgctccg gccgcgacaa cgcgttcctc tcgatccgca 19620
 acgtctccaa gcgcttcggc gaccgcgagg tgctgcggga cgtaagcctc gacgcccggc 19680
 gcggcgaggt tctcgtgctg accggtcgca gcgggtcggg caagagcacc ctgctgcacc 19740
 tggttaacca tctgcaggcg gtggacggcg gggagatcct ggtcgacggt cggcacgtgg 19800
 gctatgtcga gcgcggggga cagctcgtcg aggtcgccga cctcgcacgg gcccgcgctg 19860
 agggccggat cggcatggta ttccagcact tcgacctgtt cgagaacctg tcggtgctgg 19920
 acaacgtggc gttcgcctcg caccagctgc acgggctcag cgaggaggcg gcgcgcgagc 19980
 gggcgagaca gctgctggtc cgagtgggcc tcggcgagca cctcggctcg ttccccgcca 20040
 ccctctccg tggtcagcgc cagcgggtcg cgatagcccg cgcgctggcc gtcgaccccg 20100
 tcgtgatgct gttcgacgag cccacgtcgg cgctcgaccc cgagctggtc gacgaggtgc 20160
 tcgacacggt gcgtgggctt gcggcaagcg ggatgacgat gctggctcgtc acccatcacg 20220
 tcgagttcgc gcgggagatc gccgaccggg tggctcgtcct cgcggacggg cgcacgtgg 20280
 cccagggccg ccccgacgag gtgctgggag gcccggtcgg cgcgcgtggc tgagcagagc 20340
 cgcagcggaa tggctcttcg gttttagggt gatctgcttg tgggggtttg cgggcgggtg 20400
 agttgatcga gcgagtgacc ggcgtccggt attcgggtgac gcagacctgg acggtgttgc 20460
 gcgaacggct tggctggagc cgcaagtoga gttatcaatc ccttgccgcc gcacgttccc 20520
 ccggtcgtga cctcgtcac gctgcgacgg cagcgtcggc ccgttgcgga cgtacgtgtc 20580
 gaccgccggg ggcgagcagt actactcggg gcgctggcgg ggaagcatct tcacggacag 20640
 caccgccgct acgacggaga agacgatgct gcctatcacc actgtgctca gcccgctgct 20700
 gaacgcgtcc cgcaacagat ccgcggagat ctgcgggtgg gcgacgatcc ttccggcagc 20760

ES 2 661 447 T3

cacctcgtgc gcgatctggg ctgcgcgctc caccgggaag cccatgtcac ccagctggcc 20820
 gccgaggaca ctgtggacgt tggcggatag caaggtgccc agcgtctcga caccgacgac 20880
 caccgccgagc gggcgcgatgc tgttggcgat cccacttgcc aggccagccc gttccacggg 20940
 caccgccccc agggctgccc ccatgatcgg tgccctggcaa agcccgcctc ccgcgccgag 21000
 cagcacgaaa gcgggtccga gcagcagcca gctgggtgctg gtgggtgatca gggtcagcgg 21060
 cagcaggccc agcagcgtca gtgccaccgc cgtcgtggta acggcagatg tgctgtagcg 21120
 gctcaacaac ttgccgcccc acggcgcggt caccaggaac ggcaggatga tcggcagcag 21180
 ggcgtctccc gcagccaagg cggagtagcc ccggatctgc tggaaccaga tcccaggtta 21240
 cgtcatcagg ccgaagaagg acgccacgac cgcgaacagg atgaccagtg ccccgtgta 21300
 gttgcgctgc cggaacaggc tcaggctcgc caggggggtgg cggaggcggg gctcagccag 21360
 gatgaacacg gcgaacagga cgtggcagc gatcagtga cagagcacga cggcgtcgcc 21420
 ccagccccgg ccagccaccg agatgatgca gtacgtgatg aagaacagcg ccgccgaggc 21480
 cagcagcaga ccggtcaagt ccagcttgcg gtccggcgtt cgtcgtgctgc cggcgctgcc 21540
 ccgggaggcg atcaggaccg ccatacgatg gggcaggttg atcaggaatg cccaccgcca 21600
 gcccgccccg tcgaccagca cgcgcgcgag cagggggccc acgacgaagg aaaccgaccc 21660
 gaccgtggtg tagatgcccc gcgcgcgggc gcgttgccg ccggtgaagc ctgcggtgag 21720
 ctgcgcgagt ccgctggccg tgagcaccgc tgccgccagc ccctgcaccg cccgggcgcc 21780
 aatgagcacg cccgggggtct gtgcgatggt gatcagcatc gaagcgagag cgaaggccag 21840
 caccgccgatg ccgaacacct tgcggaagcc gatgcggtcg gccagcacgc ccgcggcgag 21900
 caggctggct ccgtagacga ggtgtgacct gtgcagcacc cattgcatgc ccgagagcga 21960
 aagcgagagg tccgtgctca ttcggggcag tgccacgttc acgatcgtcg tatcgatgcc 22020
 ggccatgccg acggccacgg ccaccacggt cagcagcctt ttctgtcctg ggctccacgg 22080
 ctgttctttc agccgcgcag cagcgctttc ctccaggac gcctctgtca ttgcggggat 22140
 tccttcagta actggattgt ttacttgtaa tcaaacctgt ctgcgcagtt atcgtcaagc 22200
 cgcgcgcaac agctcggaca atgcgctgtg caaggtgtcc gcagggccgg tgccatcgcc 22260
 cgcacacgag atatgcccg cggggcgat cagcattgog tccacaacgg acggttcogt 22320
 ggtcttgccg gcgatcaact tcaactcgatc ggaccaggct gccgcgactg agcggcaccg 22380
 cgaaccgccg ctcagctcca gcagcaccgg tctgccgtcg cgcagcagct ccgccaggcg 22440
 ggtggcggtc gcgtcggtag tcaggggaac gtcgtggcg aattccccgg ccagcgatg 22500
 ggagtcgagc ccagcggggt gacggatcca ggtgccggtg atcatctcgg tgagatatcg 22560
 gttgacctcg tccatctcga tcatttccgc gaacagctcg cgtagcgcgt cgacctgtgc 22620
 gcccgacgc atgagtgcga cctgggcgcg ggtgttggtc aatacgcggg ccgccactgg 22680
 gtgttgctcg ctgtggtatg tgtccagcag cgtcggcgac gccagccgc gcacctcggc 22740

ES 2 661 447 T3

agccaacttc cagcccaggt tgaacgcgtc ctgcaggccc aggttcaggc cctgccctcc 22800
 ggtggggaag tgcacgtgcg cggcatcccc ggcgagcagg atccgccgcg tccggtattg 22860
 cgccgccttg cgggtgttgt cgggtgaaccg gctcaggtgc tccggggtgg acagcgggaa 22920
 gtcttgccca gtgacgcgtc gcacgcctcg cctgatctcg tcgctggtcg ccgggctgtc 22980
 gaagtgcagc tcttggtccc attcgcaggc catcaccggg ttgccgagcc agttcaggca 23040
 cacaccgcga tcggttcgca cccgcccggg cgggaagtcc tcggtcggcg ccggggcgaa 23100
 gaaacccaac cgcgccacca cgggtgtccc ggatccgggg aagtcgatgt cggcctgcct 23160
 gcgcacagcg ctgcggccac cgtcgcagcc gaccaggaac ggcgcccgcg cccggtaggt 23220
 tccgtcgggg gcggagatgt cggcgacgac gccgttctcg tcgtgctcca tcgacgtgaa 23280
 ctctgccccg tggcggatgg acgcgcccg gtgcaggcg tgctcgcgga ggatttcgtt 23340
 gatctgtcc cggctgaacg ccagggcccg cggctgctcg gtgcgcacct cggacttgcc 23400
 gatcacgaac atgccggaga agtgcccgcg cagcgcgga tgcccgctcg cgtcctggaa 23460
 gtactcctgc agttcgtggg agctcgcgc cgcgtcgtcg aaggcctcca gcagaccgcg 23520
 gcgatcgagg agttcgaggc ttcgggagtg cagtcaccag atccgaggct cggctctgctc 23580
 cggtcggcac cgctcgacga caagcgtgcg cagctcgtgc agccgcaact cgcaggcgaa 23640
 caggaggccc accgggccac cgccgacgat tatcacctcg aagtctgctt ctggccgcat 23700
 ggttctccta atgcaggatt gacacggcgg cagcccgtag acgaaatgag gggattcaga 23760
 cgaacgcttc aaaccagga gcgccgcatg cggccgtag gagtaaatct cttcctgggg 23820
 gtttgtttac atggtaacat cgtcttacgg aaacggctag cggtaagtgg ggaaccgcg 23880
 ctccgggaag ttggtcccg atttcggaac attccgttg ttgcccttg tgttggttga 23940
 acacctggcg aactcgccga tgctgcctcg caaatagata tgctgtgag ggatagatgg 24000
 cagatcaggg atatccgcga tctcgtgccg aagcggcggc ggtgttcgaa gcggtcgagg 24060
 acagcgtcgc ggccctgacg cgttcgacca cgtgtgccgc gtggacgcgg gaccgggatg 24120
 atctgttgct ggtgatcgag gagccgaag gtggtgctcc aggggacttc ggcgcgctcg 24180
 ctcaagaggt gcgcggggcc gtgtcgggc agcacgacct cgtgcctcg gagatcgtgt 24240
 tcgtcccagc tcgatccatt cctcgcggcg gcaacgacga ggtgctgcgc gcggagacgg 24300
 ctgccattca tctccggggt ggcttgagcg tgctgcacgc gaccggcag ctgggagtcg 24360
 cggcgccgcc gctgctcggc gcggaagaat ccattgtgga cgacgaggtg ctgcgcggcc 24420
 tggctgacgc ggccgaaagc ggtggcgagg ccgcgccgt ggacctgaa ctcatgcggc 24480
 tggacgtcgc cgagttgctc ggtgaatcag cggagagcat cgccgacgac gaagacctga 24540
 tcatcggggg cttggactcc atcatgatca tgacctgggt caaccagtgg cgcgccgacg 24600
 ggggtgcaggc cagcttcccg gagctgttcc aggacgcggg tctggccaag tgggggaga 24660

ES 2 661 447 T3

gcatcagcgg gaccagcgag cccgaggccg gcgacgaggc ggcggtcgac gaggcggtgg 24720
 tcgacgaggc ggcgcggttc gacctgtcgg tgatgcagca tgcttactgg gtcggacgcg 24780
 gcgacgagca ggtgctcggc ggggtcggcg cgcacttcta caacgagttc agcggtcgcg 24840
 acgtggatcc gccgcggctg gagcgcgcgg tcggggcgct gaccgaacgc cacggcatgc 24900
 tgcgagcccg gttcaccgat gacggcaggc aatggatcgc cgccacctgc ccttgccccg 24960
 gtctgtcggg gcacgatctg cgtcagctcg ccgaggacga ggtcgaccac cgctgagct 25020
 cgctgcgcga gcgcctgtcc caccgcccgc tggaggtcga gcggggggag gtgttcgacg 25080
 tccagctctc cctgttgccg ggcggtgcca gccgggtgca cgtcaacatc gacatgctgg 25140
 tcgcccagcg gttgagcttt cgcacacctc tcaccgacct ggcaggtgc tacgaggaac 25200
 cggaggccga gcagccgccc atcggctaca gctacccgaa gtatctggcg cagcgcaacc 25260
 gccggcggac ccaggcggcc gacgagcgcg gcgcatactg gcgggaccgg atcgccgacc 25320
 tgccaggcgg gccggagctg ccgttgcca tcgcaccgca acggatcagc cagcgccggg 25380
 tgaccaggcg ttaccactgg ctgcagcgc aggggcggga tctgctggcg gccaaaggcg 25440
 ggcagcacgg gctgaccatg ccggtggtgt tcatgaccgc ctgcgcgaa gtgctcggga 25500
 cctggagcgc cagcaagagg ttctgtctca acctgcccct gttcgaccgg gaggcggtgc 25560
 accccgacgt gttccgcctc agcggcgatt tcaccaactt gatcctgtcg gaagtggacc 25620
 tgaccgagcc gagtccttc gctgagcgc cccggcggt gcgtgcgcag ctgcagtcg 25680
 acacggccaa cgccgagtac tccggcgctc cggtcctgcg cgacctggcc agggccaacg 25740
 gcggagaacc ggtgcgcgca ccagtgggtt tcaccagcgg gctgagcctc ggcgagctgt 25800
 tcagcgagga ggtccgcgc tgcttcggca gcccgacatg gatgatctcg cagggccccc 25860
 aggtgtggct ggaccaccag gtcaccgagc acgacggggg ctgtctgtcg aactgggacg 25920
 cggagaacga gctgttcccc gacggctctc ttgacgcat gttcgccgc taccgcgtcg 25980
 tgctggactg gctggccgca gccgactggg cgcagccggt gccgccgttg ctgccggccg 26040
 agcaagcggc ggtccgggct accgtcaatg agaccgagg gccgacgtct ggctactgc 26100
 tgcacgaagg tttcttccgc actgccgaga tcgcgcgcga acggcccgcg ttgctttggg 26160
 gcgagcaggc cgttctcagc tacggagcgc tcgccgacgc cgcccgcgc attgccggct 26220
 tcctgctggg gcgcggggtg cggcccgtg agcgggtggc gatcacgtg ccgaagggtc 26280
 ccgaccaggt ggccgcggcg ctccgggtac tactggccgg ggcgacctat gtgccggtgg 26340
 ggggtggacca gccggtggcg gcgcggcgca agatctaccg ccgcgcgag gttcgccagg 26400
 tactcaccac tgccgaggag cggacggact cgacctggcc cgccgatgtc gaaccggtcg 26460
 cggtcgaggt cgccacgtcg gtggaccgcg tgaccaccca ggtgccaggg gaccgaaccg 26520
 acctgcgcgc ctacgtgatc ttacgtccg gttcgaccgg cgagccgaag ggtgtgatgg 26580
 tcagccatgc ggcggcgatc aacaccatcg ccgacctcaa cgagcgcttc gaggtcggcg 26640

ES 2 661 447 T3

ccgaggaccg ggtgctggcg atctccgcgc tggacttcga cctgtcggtc tacgacatct 26700
 tcggcctgct ctcagctggc ggggcgctgg tgctggttcc cgaggccgag cggcgcaacg 26760
 ccgcgcactg ggtggagttg gcacggcgct ggaaggtcac gatctggcag tccgtgcccg 26820
 cgttgctgga catgctggtg accgcggccg cggctgcgga ccaggtgctg gacctgcggt 26880
 tggcgctgct gggcggtgac tgggtgaccg tcgacctgct cgaccgcctg cgggcctgct 26940
 cgccgaaggg caggctcgtc ggtctcggcg ggaccacgga ggccgccatc cactccacca 27000
 tttcgagagt cgaagagatc cgggccaggt ggagctcggt gccctacggc accccgctgc 27060
 gcaacgtccg ctgccgggtg gtcgatgagc ggggccagga ctgcccgac tggggggtgg 27120
 gggagctgtg gatcggtggc cgtggtgtcg cgctgggcta ctgcggcgac cccgaccgca 27180
 ccgccgaacg gttcgtggtg cgggacggga tccgctggta ccgcaccggt gatttgggcc 27240
 ggtactggcc ggacggttcg ctggagtcc tgggtcgcgc cgaccaccag gtcaaggtgc 27300
 ggggccatcg catcgacctg ggtgaggtcg aagccgcgct caacagccac cccgacgtcg 27360
 cccgcagtgt cgtggtggtg acggtggcgg ccggtacccg gctggtggct gccgtgttc 27420
 ccggtgggcg tgtgcgcct ctggaagacc tgcgggactg gctcggcacc cagcttccc 27480
 cgcacatggt gcccgccgac ttcgtggtgt tcgacgagct accactcaac gccaacggca 27540
 agatcgaccg caagcggtc gttcgactgc tggcgccgga agagcgcacg acggcgagc 27600
 tgcggtcgcc gcaggtccg gtggaagagc tcgtggccca ggccgtgctg gatctgctcg 27660
 gcggggagcc gatcgagcgc gacgacgact tcttcacccg cggcggtgac agcctgctgg 27720
 ccaccgcgt cgtgggcagg ctgcgcgcgg ccggtctgcg cggggctgag ctgcgggcgt 27780
 tgttcgcaa gccggtgctg cgggacttcg caacagacct cgaaatcggt gcgtcaccgg 27840
 ttgcgagcgc gctcccggt gaccgggcc accggtacga gccgttccc ctcaccgacg 27900
 tgcagcaggc ttacctgctg gggcggggcc gggacttcgc gctcggcggt gtcgctccc 27960
 actggtattg ggagttcgac accgcgac tcggacctgc gcgcttgag gaggcattgga 28020
 accgggtcgt cggccgccac gagatgctgc gggctcgtgt cgaggacggc cggcagcgca 28080
 tcttgcccg ggtgcgcgc tatgccctct ccgtcgaaga cgcgcgtggc ggtgctgagc 28140
 aggtcgcgct ggaccgaatg cgggccgcca tgtcgacca ggtgctcgac gtctcgcgt 28200
 ggccgctggt cgacatccgc gccgtgcgt acagcgcgcg gcgggtgcgc ctggggttca 28260
 gcctcgacct gttgatgctg gacgcgctga gcatcgtgat cgtgttcgcc gagctgagcg 28320
 agctctacca ccgtcccgga acccggtgc ccgagatcgg ggtctccttc cgcgactacg 28380
 tgctcgggtc ttgcggggat cgagagcagc tcgccgagag caagcgctac tgggcagagc 28440
 gggctccttc gctacctccg gcaccgcaac tcccgctcg caccgatccg gaactcgtcg 28500
 cgcagcccg cttcaccgc aggcagagcc ggattcccg gcaacgctgg caacgcattg 28560

ES 2 661 447 T3

tcgaacgagc gcggcagcgc gggtctaccc cggcttcggt gctggcgact tgcttcgcgg 28620
 aggtcctggg acgctggagc aggcagcggg agctgacgct gaacctgact ctgttcgata 28680
 gccgcgaggt ccaccccgag atcaatggcg tggtcgggga ttccaccgcc ctgctgctgg 28740
 tgggttacga gccgagggcg ggggagtgtc tcgccgagcg cggggcccgg ttccagcgcc 28800
 aaatgtggca ggacctcgac caccgcgacg tctcggcagt gcaggtgatg cgggacctgg 28860
 ccgcttcccg cggatctgcg gccgctatgc cgggtggtgtt caccagcgcg ctgggcatcg 28920
 gccgtgacgt cggcgcgag cggttcgcgg aggaggtttg ggggctgtca cagacccgc 28980
 aggtctggct cgatcaccag gtgcgcgacg acgagggcgg catgcggttc aactgggacg 29040
 cggtgaggga gttgttcccg gaaggggtgc tcgacgccat gttccaagct gagtgcgcgc 29100
 tgctggactg gctggccgat gccgactggt cgcagcaatt gccagacctg cttcccactt 29160
 cgcaacgagc ggtgcgcgag caggagaact cgacccgcga accgaccgcc gcagcgacgc 29220
 tgctccacag cggattcttc gctcgtgcgg acgccgaccc ggcgcgtacc gccctggtgt 29280
 ggggtgcgga ccagcgggtc gactacggcg aactggccga cctcgccttg cgggtggccg 29340
 cgcacttggt ggagcgcggg gtgcgggcgg gcgacaatgt ggccgttgcg ttgcccaagg 29400
 gacgcgacca gatcgttgcc gtgctcggcg tcctggccgc tggagccgtg tacgtgcggg 29460
 tgggcgtcga tcagcccgat cagcgcaggc agcggatcta ctgcggggcc ggcgtgcgcg 29520
 tggtgctgga cgacctgcg gaggcgtacc gccggaagcc cttggcgggt ccggttgagg 29580
 tggatccgga gaggctcgcg tacgtgatct tcacctcggg ttccgaccgt gagccgaagg 29640
 gtgtgatggt cagccacgcg tcggcgatga acaccatcgc cgacgtcaac gcccggttcg 29700
 gactcggcgc cgaggaccgg gtgctggcga tctccgcgt ggacttcgat ctgtcgtgt 29760
 acgacatttt cggcttggtg tcggtcggcg gttcggctgt gctcatcgac gaggacggaa 29820
 ggcgggaagc ccgcgaatgg gccgggcttt gccgccgctg gaacgtcacg gtttgaaca 29880
 ccgtgcccgc tctgctggac atgctgctgg tggtagcaga cgtcctgcc gggtcgtttc 29940
 ggctggctct cgtctccggc gactgggtcg gcctggacct ctgggaccga cttcgcgagc 30000
 gggcaccggg gtgcgcgctg gtgcgcgttg gcggtgccac cgaggccggc atctggtcga 30060
 acttcttcga agtcgaacgg gtcgaccgcg gatggcgctt gatccctac ggcagacgc 30120
 tggccaacca gcggttcgcg gtggtcgacc ccaccggtca ggactgtcct gagtggaccg 30180
 agggggagct gtggtcgtg ggcgcgggtg tcgcggccgg ctactgcggc gacgccgatc 30240
 gcaccgccga gcggttcgtc gaatgggacg ggtaccgctg gtaccggacc ggggacctcg 30300
 ggcggtactg gcctgatggc gtgctggagt tcctcggccg ggtcgacaac caggtgaagg 30360
 tgcgaggtca ccgcatcgaa ctgggcgaga tcgagggcgg gttgaccgcg catcccagc 30420
 tcggccaggc cgtcgcgcgc tccgtcgggg accgggctcc cgcgctcgtg gcggtggtgg 30480
 tacccgaggg cgaccggtcc tgcgattccg acgaactgcg atcgtggctg ctccgaccgc 30540

ES 2 661 447 T3

tgcccgcccta catggtgccc gagcggatcg ccgagatcgc tgaggtgccg ttgaccgcca 30600
 acggcaaaact cgaccgcgcc tcggccatcc gcgggctggc cgacggcgcg cggaaccgg 30660
 ctgagaccgg tgaaccgctt cgcggcgagg tggagatcgc gctggccgag caatggcgcg 30720
 aagtgtctgg cgtgcggcg ctcgccctg acgacaattt ctctgtactc ggcggcgaca 30780
 gcctgttggc caccgcgctg gtggagcgca ttgcgcgcct gtccggcggtg gacatcgctg 30840
 tgcgcgagtt cctcgcgggg ccatcgcttc agcagcaggg ggaattgtgtg aaacagcggt 30900
 cgcaggccgg tcaccagatg gaggaaggcg tcatctgatg gcccatcgcg ctacctcgac 30960
 cgatgggtgc ggatggggta cgctgcctgt ggaatacgcg tatacaaaac tctgtgtcgc 31020
 ggagtctctg atgcagaccg agcgcaacaa actgctttac cggcgctgga tcgaagagat 31080
 gtggcacgcc gatcccgca acatggatca gctggcgcc gagctggcg caccgggtt 31140
 cgtcgcgcac tggcccgacc gggaagcgcg cggtcgggag gcgctggccg aaatgggtccg 31200
 acggggggtc tcgatgttca ccgacgtcaa ggtggccgtg gaggtgggtc cgtggcgga 31260
 cggcgacctg gtgtcggggc gctgggtgtt tcgcggttc tacaacggcg aactgccggg 31320
 ggcgccagtg gcggcgggta ccgaagtgtt gtccggggt atcgacgtga tgcgtgcgc 31380
 cgacggcaag ttctcgtcgt actgggtcag ttccgatgtt ttggaattca tgaacaact 31440
 cgggttcacg ccgaagccct gactaaatcc aggtgttctt tgtcgccact tccgacgaac 31500
 tgtcgggaagt ggtcggcggtt ggttcgcgtt aattgtgtga acgcggtgc taatagtgtg 31560
 aacgggtgct gctaagtcct ccgctgctgg attttgatg ctgtcgtcg tggttcggcg 31620
 gtgctatgtt ggcgacagcc gaagatctga attcagcggc gaatgcaggt gaagctgggg 31680
 gggctggaac accatttccg gcttcgtgga aatctttctc cagttcgtc gctgctcgtc 31740
 gaatcgtcgt tttttgggtg ttcacgacgg tgggtgcgat tcggtaagca cagcgcgcg 31800
 tccgacctgg gaggatgaca gatgaatgcc gcagagctgg tcgcggtgct cgaagggtg 31860
 ggggtgagcc tgtggcgga aggcggtcag ctctgtttcc gggcgccga ggcgtgctg 31920
 actgacctcc accggcagca gttgcgtgcg gataaggaag cagtcctcga cgtgttcgt 31980
 gcccgggccg ggcaggtgag cgtcgtcgcc gatggcggtg cgcggttcga gccgttcccc 32040
 gtcaccgaca tgcaggccgc ctacctctc ggccgcgcg acgtcttctc ctacggcggg 32100
 gtgtcctgcc acggtacgg ggagctcgag ttaccgagc tcgaccgggt ccggttgag 32160
 gccgcctggc agcgctgat cgagcgccac gacatgctcc gcgctgtcgt ggacgacgac 32220
 ggtaccagc gggctcctgcc aaaaacctcg ccctaccgca tcagtttcga cgatctttcc 32280
 gaattgggcg agggagcgcg ggcgaagcgc atcgccggga tccgggagga gatggaccac 32340
 cgggttcagc ccaacgacgt gtggcgctg ttcgagctgc gggtcagcag ggagccggcc 32400
 ctgtccggc tgcacttctc catcgacttc ctcatcgcg acttcgtcag cattcagctg 32460

ES 2 661 447 T3

ctgctcgaag agctccggca ctgctacgaa cccccgacg aaccgttgcc gccgttgag 32520
 gtgacgttcc gggactacct cgcgcagaa cgccaactgc gcagcggccc ccgctacgaa 32580
 cgcgaccgcc agtactggtt gggccgcacg gacgagctgc cgcgcgccc ggatttgccg 32640
 atggccgaaa ccaccggttc cgtggccag gcccgtttcc ggcgctacgc gttaatgctc 32700
 gaacccgagc agtggtcggc gctgcgggat cgcgcaggg cgaaggcat gacgccatcg 32760
 acggcgggtgc tggccgcata cgcgaagtc atcgggaagt ggtgccgccg cccggacttc 32820
 acgctcaacc tgaccttgct caaccgacag ccgctgcacg ccgacgtgcc ccgcatcatc 32880
 ggcgacttca cctcggtcga cctgctcggc gtgtggcaga accgtgcgac gagcttcgct 32940
 gatcggggcgg cgcgctgca ggaccaggtc tggcgcgacc ttgaccaccg cctgttctcc 33000
 ggcgtcgagg tcatgcgcga aatcgccgc cgcacgact ccgacgccgc gctgatgccg 33060
 atcgtcttca ccagtgccat cgcgctcggc ggaaccgacg atgggctcgg cgaaaaccag 33120
 tggggcaggc tcggtacgg catcagccag accccgagc tcttcatcga ctgccagaac 33180
 atcgagcgt caggcggcct ttcgacgaac tgggacgtgc gggaaggcgt gctgcccgag 33240
 tccgtcgtcg cggacatgtt cgcggcatat cgtcagctgc tgaccgcgat gtgcaccgag 33300
 gacgaggtgt ggaccaagcg gttcccggtc gacctgccgg agcaccagcg agttcgacgg 33360
 cactcggtag acgacaccgc ggcgccgatg ccggacgcgc tgcgcacga cggcttcac 33420
 gcgcaggccg agcgcgcctc gcaccggacc gccatcctca gcgcccgcg cgagctgacc 33480
 tacggcgagc tgctcgaccg cgcgcgcggg gtggccgccg cactgcggga agcgggctgc 33540
 cagcgcgggg acgtcgtggg cgtggtcatg gacaagggtt gggagcaggt ggccgcggtc 33600
 ctgggcatcc acctcgggtg tgcgcatac gtcccgggtg aactgcacca accggaggtg 33660
 cgcgctgagc ggatgctcgg cgatgccggc gtgcggctgg tgctcaccca aacttggtg 33720
 cagcagagc tgggcctgcc cgagggcac acgagcatcg tcgtcgacga gcacgggccg 33780
 gctggagcct ccgcagcgc ggacgagcgg gccgcgccc acgacctggc gtatgtcatc 33840
 ttcacgtcgg gctcgaccgg ccggcccaag ggcgtcatga tcaccacccg cagcgcggtc 33900
 aacacgatca acgacatcaa caccagattc ggcgtgaccg gcgacgaccg ggcgctgatg 33960
 ctggccaatc tcggcttcga cctgtcggtc tacgacatct tcggcacctt ggccctcggc 34020
 ggggcgctgg tgatccccga gcacgatcga cgatccgac cctcgactg ggcggagctg 34080
 gtcgccgagc acggggtgac ggtgtggaat tccgtaccag cgcagatgca gatgctccag 34140
 cagtatctgg acgccaaca gctccggctg ccgcgcgcc ggcgcgctc gctgcgcctg 34200
 gcgctgctct ccggtgactg gattccggtc gggctgccc cgcagtcga caagcacctg 34260
 cccgagctgc ggatgatcag cctcgggtgt gcgaccgaag ccgccatctg gtcgatcttc 34320
 cagcaggtaa acacgtcggg gatccaccag cgcagcatcc cctacggcaa gccctgacg 34380
 aaccagacct ttcacgtgct ggaccaggac atgcaggact gtccggaagg gacggtcggc 34440

gagctctaca tcggcggtgc cggggtagcc ctgggatacc tcggggatcc agagcgcaact 34500
 gccgagcggt tcgtcgtcca cccaccagc ggctgcgtc tgtaccgcac ggtgatctg 34560
 ggccgctacc gcgaggacag tgagatcgag ttctctcgcc gcgaggactc gcaggtcaag 34620
 atccgcgggc atcgcatcga attggtgag gtgcaggctg cgttgcagtc ccaccgcag 34680
 gtgggcccgg cgatcgcggt ggtcgacggg gagcgctcga cgaacaaacg gctcgcgcg 34740
 ttctagaaag gccgcgagcc ggaggcggag ccgccgaccg cggagccggt ggtgcgggag 34800
 gtgaccgcgc ccgcgcagct gatccggcgc ggtttggaag ccgacattga tcgcgaccgc 34860
 atcgcgaggt tctctgagcg gctcgacgcg gccctgctac acgccatgac gcacgcgctc 34920
 gccgacgggt acccgttcgc cgggtggagcc gtcctttccg ccgaccaggt cgcgcgcgcg 34980
 ctcgcgctcg cgcaggtgta ccggtcgctg gtgccgcggt ggctcgccga actggcgggg 35040
 gccgggttcg tcgagcatga cgagaacgga taccacgacc tgcgcgtggt cgaagccgaa 35100
 caggtcgaac ggtgctgggc cgaggccgag gcggtgcgac cggagcggat ctgcccgacc 35160
 gcgctgatgg gctacttccg cgactcgcg cggaactgc ggcaactgct gcgcggcgag 35220
 gtcagcgctt cccggtggt cttcccgaa ggccggtggt acatcgcgga ggcgacctac 35280
 agcgacggcg tggcagcaga ccacaacaac caggtcgccg ccagggcgct gcgcgcgac 35340
 gctgcgcaaa ccggcgggaa acggctcgcg gtgctggagg tcggcggcg catcgccagc 35400
 accaccggc acgtgatgga cgtcctcgcc ggtttcgag tcgactaac cttcaccgac 35460
 gtgtcccagt tcttcctcaa cggcgcgcg agcgcttcg ccgcgcagtc gatggacatc 35520
 cggatctacg acatcaacca ggactaccgc gaacagggcc tgcgcccga gtccttcgat 35580
 gtggtgctct gcggcgcggt gctcaacaac gcccgcgata tcgctggac cctgcgccag 35640
 ctcaaggaga tgctggtccc cggcggttg ctggtgttca tcgaggcgac tggtagcac 35700
 ttggagatca acgtgtcgca ggcttcatg atggaggagt tcgaggactt ccgccgcgag 35760
 gagggcacca cgttcctcga ccggcgcgcc tggctggaac tcgtcggtc cttggatgcg 35820
 gactgggtga cctgtctgcc cgcgcgaaac gaccgcctct cctgctcgg acagcatgct 35880
 ttctggtgct gcctgcggcc cgggaagggt cgggtcgaag ccgccgacgt cgtcgagcac 35940
 gcccgcaagc tgctccccga acacatgctg ccggtgaaa tgcaggctgt gcacgacatc 36000
 ccgctcagcg acaacggcaa ggtcgaccgc ggctgctgc agcagtggtt ggcgcgcagc 36060
 gcgccaaccg aacgggcgga caccgcgag gaccacggtg acgagctgga ggagcggctg 36120
 gcgagcatgg ccgccgacgt gctcgcgctg ccagcgctcg gccgagagca gaacttcttc 36180
 gactcggtg gcgattccct gctggtcgcc cagctcgtcg gccggatacg ggagcaggtt 36240
 cccgaagccg ccggcctgta cttcgacggc ctgctcgccc gcatacctcaa ccagccgaca 36300
 attcgggcgc tagccgagca catccggcag acccagcgga ccgagcacga ggcggtggcc 36360

ES 2 661 447 T3

aataccgcga tgcgccatt gatcgtgctc aacgatccgg gtgcggagcc gctgcggggtt 36420
ctcgtgcacg aagggatcgg cacgatggcg ccgtaccggt ggctggccaa ggaactcctc 36480
gaggcgggac cgctggtcgg gctcgcggtc aacgacacgt ccggttacct gggcctgccc 36540
gccgaggagc tgatcccgcg actcgcggag cgatacggcc gtgagctgct ggaccagggc 36600
ggtgagcagt tccacatcgt cggctactgc ctcggcgggc tgctggcgac cgagatcgcc 36660
ggtcggctca cggaggccgg tgcgcacgtg gcgagcctgg tcgtggtcag cagctaccgc 36720
gtgccctacc agattggcga ggagctgctg ttcgagtacg ccttcggccg ggtgctgtcg 36780
gccgatctcg ggaggatggg gttccccacc gacgaacagc tcgtggggca ggcgttccaa 36840
cgagtgtctg cggcgacccc cggccaagtg ccggacgggg cggttcgacgc gctcgacggg 36900
cagccaggcc tggtggaagt ggcggagcgg ttccgggcgc tgcgccgacg ttcaccggaa 36960
gaacgccggg ccgcgatggc gcggcagttc gttgatgagg gcctcgcctt cggctccccg 37020
gaacaggcga ccagctgta cgaggtgttc cggcagagct tccgcgcggc gaacctgcat 37080
cagccggagc tctacgcggg tgacatgacg cttctgcggc ctgggggcaac gctccggttt 37140
ctgcccggcc tgcaggcggg catgaccgag ttctggcgcg atgtctgcct cggcgagctc 37200
gacgtgatcg acgtcgcggc cgaccacttc agctgcctgg acgcgtcgaa cgtcgcggaa 37260
gctgtccgcc acatccgcgc ggtggcgag aagtcccgag tcagctgatc tccagtcgga 37320
agaaggagca catgagcaga gtccctgacc ccgggcgcag gccgaaggtc gtggtgtgcg 37380
ggaccaggtg gggccgcgtc tacctccagg cgttccctgga gccggacttc ccgttcgaac 37440
tcgccgggat cgtcgcgccg ggcagcgacc ggtcccatgc ctgcgcagag ttctaccagg 37500
ttccgctgta caccaacgtc gaccagctcc ccgacgacat cgacatcgcc tgcgtcgta 37560
tcagcgcagg agtcagggtt gggcgaggcg cggaaatggc gcaggccatc atgggcccgcg 37620
gcatccacgt gctgcaggag cagcgttgc accacgatga gctggctgaa tgcctgcggg 37680
gcgcccggaa acacgggggt cagtaccggt tgaacaccca ctaccgcac atcgagccgg 37740
tccgcgggtt catcgccgca gcacgggagc tggtcagcaa gcagcggccg ctgttcacatg 37800
acacgacctg cggcatccaa gtgcgctaca cgctgttcga catcctgggc cgggcgctcg 37860
gcgagggtccg ccggtggggg ttgcgacgac tgcgggtggt gtccgaccgg attcgcggga 37920
tgaccgacct ggagttcccg tacgcgctcg tgaacggcgt gtccgcgggt gtccgatgg 37980
tgctgcgtgt gcagaatcag ctcaagtcca gcgaccgga caactacacc cacctcttcc 38040
accggtacac ggtgggcacc gacggcgggc acctgacgtt ggccaacacg cagggtccgc 38100
tgctctacag ccccgcccg caggccccg agcgggcgtg gcacacggtg cggctgttcg 38160
acaccgacgc cgagcatctc acctaccgga gcgccaccgt catcggtgac gtcgaaggac 38220
ccgactaccg gcgattgtc gccgatatct gccaggtgg tgtccgcaca gcgctaggtg 38280
cgctacgcca tgcgcgctc tccggcaacg atccgctcaa gcaggggcag tactacctga 38340

ES 2 661 447 T3

ccctgtgccg gctgtggcag gacctacta ccaggctcgg ctatcccgat cagctcacc 38400
 gcacggatcc gccggacgtg ctgccggtcg acgacatgat ggccgcccg cttgcagtcg 38460
 aagtgaagga gggctggggg tgactagcac gatggcgccg ttgtaccagc gctggttccg 38520
 gtgttatcgg cccgtggagc actgcgggg ccggctgata tgccctgcgc acgccgggtg 38580
 ttccggcagc gcttttcgca cctgggcgcg cgccctgcg ccggacgtgg aactgctcgt 38640
 cgcgcagtac ccgggcaggc aggaccgcat ctccgaaccg cacgtccggg acatggacgc 38700
 gctgttgcc ggtctgatca gggcgatgcg acccgtgctg gaccggccgt tggccttgct 38760
 gggcaacagc atggggggcg cagtggccta cgaactcgcg cagcatctcg ggtcggcgtg 38820
 cgcgcagctg ttctgttccg ggccggcctg ccgagccga cagcgaccg ggggtgtgca 38880
 cctcagcggg gacgacgggc tgctggccga cgtgcgcaat ctccggagtg tcacaccgca 38940
 agtgctcgaa gaccgggtgc tccgcgaact ctctctgcg actctgcgcg ccgactacca 39000
 gctgatcgag acttaccggc cgacgacaag cccgttgccg gtgccggtga tggcgctcat 39060
 gggcagcagc gaccccgagg tcgacgaggc cgatgccgcg gcctgggcgg cggccacgac 39120
 aggcgagttc gagctgcgct gggtcgacgg tggtcacttc ttcttgaccg accgaccgga 39180
 cgacgtgggt aaggtggctg ccgagcgccg ggccggatcg ttccggagtgc gcgatggccg 39240
 ctgaatccga ggctcacgca cgcgggttcg cgatccgcat gttcaccgcc gcagtcgccg 39300
 cgcaggagct gttcacaagc tacctcgccg tgaagctcgg ctctaccgg gagttggccg 39360
 ccggtggacc ggcgaccccg ccggagctgg cccggcaatg cggcgccgac gtccggtagc 39420
 tgcgggaatg gcttagcag caggccgtgg cggaatcct ggaggtgcac gaccaggcaa 39480
 tgcccgaact ggaccgccgc tacctgctgc ccgcaggaca cgcggaggcg ctgctgggta 39540
 ctgatagccc gtactcgatc tcgccgatcg cgggtgctgc gatgggcgcc accggccggg 39600
 tgctgccgcg gctgctggac gcctatcgca ccggcggcgg cgtgccctac accgaatacg 39660
 gcgaggattc ccgggacggc cagcccggtc ccaaccgagc gctgttcacc aacaacctgg 39720
 ccgactggat tcggcgtttc cttccggaca tccacaaccg gctgctcatc agcggccgtg 39780
 tcgccgacgt cggatgtggg ggggggtggg cgagcatcgc gctcgcgaa gcctaccccg 39840
 atgtccacgt cgacggcttc gatctcgacc agcagtcctt tttggatgcc cggcgcaacg 39900
 ctgcggagca cggcgtggcc gaccgggtca gcttcgaggt ccgcgactgc ggcgatgccg 39960
 aactggccgg tagctaccag ttggtttgca tcttcgacgc gctgcacgac atgtcgcgtc 40020
 cggtcgaggt actgcggacc tgccgtgtgt tgccgcgcaa cgcgggcacg gtgctggtga 40080
 tggacgcaa ggtgcccgat cagttcaccg cggccggcg ggacgtcgag cgcttcctgc 40140
 acgcggcgag cgttctgcac tgtcttcggg tcggcaggtc agaagagcct tcggcgggca 40200
 ccggaacggc gatgcggccc tcgaccgtcc ggacctacgc ggtgaaggcg ggtttctccg 40260

ES 2 661 447 T3

acatcgcggt gctcccggtg gacgacctgt accaccgctt gtaccggtg ctcgactgaa 40320
 gtgccgaaat cggactgggg aggaaagggt accggacgac tcgcacccgg aagcagctcg 40380
 tgccggtgat ctctgcatgg cgagggccgg cttagacgct tgcaaccgcc ccggcccttg 40440
 gcggccgcag tggtaaaggc cagggccgaa ctgcgccggg aatgcgtcaa cgccgagcag 40500
 gcgcaccggc tcgtgcaccg ctctgggttc ctgctgtagc cgcagattgc gctcgtact 40560
 aggggtttcg cggaacagtg ccgggaaagg ttgatcatgc gcacgctgat tatcgacaac 40620
 tacgattcct tcacgtacaa cctcgctcac tacctgcaag acgtgaacgg cgaaccgtgc 40680
 gaggtcgtcc gcaacgacga tccgacgttc cggttgagt atctcgaccg gttcgacaac 40740
 gtcgtcctct ctctggacc tggccaccgg gaacgccggg cggacttcgg ccaactgcgcg 40800
 gaggtcatcc ggcacgcggc cctgccggtg ctgggcgttt gcctcgggca ccagggttg 40860
 tgcctcgctt acggcggaac cgtggcgggg gctccggagg tccggcacgg ggggctctcg 40920
 caggtgaacc acgaaggccg gggcctgttc aacggactgc cctcgccgtt ccggccgctc 40980
 cgctaccact cgctggtggt gaccagctt cccgacgagt tcgaggtcga cgcgtggacc 41040
 ccggacggcg tgctgatggc cgcacatcac cgaaatcgtc cacagtgggg agtccagttc 41100
 caccggaggt cgatctgcac cgagcatggc cacctgctgc tgtgcaactt cgcgagctg 41160
 accggccgct ggcagcggga aaacgcacgg ccagcgttac ccctgcgggc ccgggtcgcg 41220
 gcgacaccgc acgtgcgca ggtcgaggcg aagctgcggg tgcctggcga ggaactaccg 41280
 ggatggtgcc cgccagagcg ggtgttcgac gcgctctacc aggaaagccc gatttcgttc 41340
 tggctggaca gcagcctcg cgagggggag cagggccggt tctccttcat gggtagcgcg 41400
 tccgggccgc tggcgcggtt cgtcgaggcc gacgtctggc agggcttggc gatcgttcgg 41460
 gagggggacc ggaccattga ggtgcgcggc ccgttcctgg actggctgga gtccgatctg 41520
 gcggcgctgt gggccgatgt gcccacctg ccgttcgatt tcgcgctggg ctgggtcggc 41580
 tacctcggat acgagatgaa ggccgaatgc ggcgcgcggc gggcgacccg ggcgcagacg 41640
 cccgacgcgg gcatgatctt cgcgcaccgc gccgttcgct tcgaccacca ggagaagcgc 41700
 gcctacctgc tcgccctggc cctggtcgac tgcgagcagc ccgcgcagca gtggttggtg 41760
 gagaccggcc gacgatggcg ggaactgcgc tctcaacagt ccgatgagga cgttcgggca 41820
 ccggccggac cgttcaggga gctgcggctg cgccacgac ggcacgacta catcgactc 41880
 atcgaacggt gccaggagca gatccaggcc ggcgagagct acgaggtctg cctgaccaac 41940
 atggccaccg gcgaatgcga cctggatccg tggcacagct acaagtctc gcgcgcgcac 42000
 agcccggcgc cgttctcagc gctgttgagg ttccggggcg tgtcgtgct gagcacctcg 42060
 ccggagcggt tcctgcggat ttcccgggac cgggcggtgg agtccaagcc gatcaaggga 42120
 acgcgacccc gctcggacga cccgtggag gacgcacgcc tgcgcgagga actgcggacc 42180
 agcgagaagg accgcgcgga gaacctgatg atcgtcgacc tggtcgaaa cgacctgggc 42240

ES 2 661 447 T3

aggtgtgctg aagccggttc tgtccacgtc ccgaagctgt tcgacatcga atccttcgcc 42300
 accgtgcatc agatggtcag caccgtgctg gcacacctgc ggccggacgt ctcccccatt 42360
 gctgtgctgc gcgcggcctt ccccgccggt tccatgacgg gtgccccgaa gatccgcacc 42420
 atgcagatca tcgacgagtt cgaaagcggg gcgcggggta tctactccgg tgcactcggc 42480
 tacttctccc tgagcggggc gaccgatcac agcatcgtca tccgcacgct cgcggtgaac 42540
 gctgggcagg tgagcttcgg ggtcggcggg gccatcactg ccctgtcggg tcccgaagag 42600
 gagtggaccg agacggcggg gaaggccgct ggactgctgc ggttggttcg cgcgggcttc 42660
 ccgggcaatg acgtggtcgc gaccgaatga gtccgcatgc aggtcctatg tggacgaaag 42720
 ccaaccagtg atcatggctt tcaggccctc ggggtagctg tcgggtttcac ggcgggtgga 42780
 ccactggtcg gagaggcagg cgaggttggg catctgctcc cggtcgacgt cggagaccag 42840
 cacgatcgcg aaggcggggc ggtcgagtcc gtccgtgctg cggcgggtgg cagtgcgcac 42900
 gagcgcttcg ccgagagtga agtgccacac cgcgcggtag gccgcggcgg cgcacttgctg 42960
 gtcgagtcgc gccttgacga acgtggcgag gatttcttcg acgatccaga gcaccgaagg 43020
 cgccatcatg tcgcctgctg cgagtaactc gatcggccag ggatggcggg tgagctgtcc 43080
 gtgcaggagt ttccagagct cgatgaggcg atcttggggg ttgtcgggca gttccgggcg 43140
 gggtaggcgg gcgacgagcc tgtccaggat gaggacgagg agttcgtcct tgtcctggac 43200
 gtggcggtaa acggccatcg gggagctgcc gaggttcttc gcgaccgggc gcatcgtgag 43260
 ggcttctatg ccacgcggga cgaggatgct ctccgcgacg tcgacgatgg cgttctggga 43320
 cagccgcggg gggcggccgc cccgattcag agtcttgcca gccgggctca tcctgccatt 43380
 atgtcagccg gcgggtgcag ccgggaccga actcaccacg tcccacttcg caagatgaag 43440
 cggcgagact tgccggtgcc ggtgcgcggg agggcgggtc cgaccgcgac ggaccgggga 43500
 accttgaacg gcgccagccg gtcgcgcact aagtcgagca attctcttcc cagatcctgc 43560
 acttcgcgct gccgggtggg gacgatgtag gcgcgcagct tggtcgcacc ccgctggtcg 43620
 gcaatcgcgg ccaccgcggg ttcttgacag gtggggtgct cgtcgaggat ccgttcgatt 43680
 tccaccggat ccaccgtgat cccgccaacc atttccagat cgtccttgctg tcccttggtg 43740
 atcaggcgcc cccggtcatc gaatgccgca cggtcgcggg tgttcaacca tccgtcgcgc 43800
 agcgctgctg cgggtcggttc cggctggttg aggtagccgg cgagcagcgt cggcccggtc 43860
 acccagagct cgccgtcgac accgcagtca acgatgtccc gaccgctgtc gtcgcgcagg 43920
 cgcacctggt aaccggcag cgcaagccg agcgtgcccg gtcctggtga gtcgagcgtg 43980
 ttggcgcaga acgcgcggcc aacctcggtg cagccgagct cgtcgagcac gggggccccg 44040
 aggatgtcgg tcaaccgctt cccagctcg gcagcgagcc gttcaccggc cgacaccgca 44100
 acgcgcagcg acgggaaacg ctgccgctcg atctcggcga ggaggttggc gtatgcggac 44160

ES 2 661 447 T3

ggcaccgcgt gcagcaccga gaccaggtgc cgatcgatcg ctccggccac ggtggacggg 44220
 gccgggtgct ccggcagcag caccgcggcc gatccggaca gcagcgggta aacgaacgag 44280
 ttgccgaagc cgtaggcgaa gtggagcttc gacaccgaca agctgacatc ggcttcggtc 44340
 atgtgcagca tcgccgaacc gacagcgcgg tggtagtgct cgagatcggc gtggcggtg 44400
 ggaacgccaa tgggacggcc ggtgggtccc gaggtgtact gcacgtacag cggggcgtgc 44460
 gccggtaccg ctggggcggc cgcgaggggg gaggtttcgg cgcgggcgag aagttctccg 44520
 gcgtcgaggc accggacgcc atcgaatcgc tccggcgagt cggtagtgac gagctgcgcg 44580
 cggcagtcct cggcgatgaa ggtgtggtcg acgaccgtca gcctgggggt ggtggtcacc 44640
 gcgacggcgc ccaggcgtgc gagggcgagg aaggtcgcaa tccaggcgat ccggtccggc 44700
 atgctcagca gcacgtgctg gcggggccgg acgcccgcgt cggccagcac ggttgccggt 44760
 cgcgcggcga ggcggtggac ctggccgtgc gtccagctcg aagagccgac gaggttaagcg 44820
 ggggttgctg gccagttgtt gcgcgctgag cggtcggcga gtgcctgcgc aacgttcacc 44880
 gtttctcctt gtgcagcagc agcggatcag cgggcgggga ggaactcgat gagctggtag 44940
 cggatcaccc cgcgggatgg cgaactcgtg ccgacggcac gccggacggc cgggggtccac 45000
 ccgtgggcgt cggccaaccg cgcgactcgc ctgccgtcgc cgaggctggt gaaaccgagc 45060
 agcaacctgc caccgggtt cagccgttcg cccgcttggc gcagaaaagt ttcgtgcgcc 45120
 tggtagtcga tgtcgtagaa agctttttcc agatcgaaa tggcgatggt ttcggcggtc 45180
 gtcggcacga agttcgagtt ccataaaatg gtgtcatatc gttctcccg gggaagcggg 45240
 gagaacatgt cgcgcagtg cgcggtgagt cggtcggcca ccgcgtgccg tctggcgttc 45300
 agctcgatgt tgtcggtgcc ggctttgttg atgtcggtg gcggaaccag gcggcagccg 45360
 cgcagcgcgg cagcaccgc caggtatccg gtgcccgcgc cgatctcgca gaaacagcca 45420
 ccggtcgggt agcggaccca ctccgcgtaa agggctgctg agcgggtcag atggggtgcg 45480
 tacacgcctg ggagcaggtc ccagtcagg ccgagcaagc ggaatgtggt tgcgttcgcg 45540
 atctggcggg gggcttcgat caggcgggtg accatttctg cgtggaggc gtccatgacc 45600
 gcgtgttgct cctctgcgaa acagtcggat tcagggcgag tctaccgatt ccgggtcaat 45660
 ggacaagggc tagaaaattg ccaaataaaa gagtgattta gatcacatat cgccctttgt 45720
 tggggtttgt taacgtgtac gcgtgaaggc aagtgcggg tatccgagga attgccggca 45780
 gggcggttcg gttgcgcttg acgacctgtg cgaaaatccg ttcgacgagt tctcgagatg 45840
 gcattccgaa ctgggcatga agcgtcgccg ggggcatttc cgcgccgtac tggcggccac 45900
 cgggctggac ggcgcgccag aagtgcggtc cgtcgcggtg atccgggtga cgggcgggtt 45960
 cgcgttcttc gccgcccacc acggagcggg catcgccggg ttccggggcag acccccgtgc 46020
 cgcggtgtgc ttcgactggt cggagcaccg ggccgaggtc agagcctgcg gccaggttca 46080
 gcgcttgctc gatgtggact ccgatggtta ctccgggacg ctgccggagc gagagcagct 46140

ES 2 661 447 T3

gacgatctgg ttgcgtgacc aggaaccgac gagcgcgcgc cgagccgaga tgacgcgcgc 46200
gctcgcgcgc gtcgcggacc acttcggggg gcagctggtt ccccgccccc agcactgggg 46260
cggttaccgg atcatcgtcg atgacatgct tcttcgcctg cggggcgcgc acgaggtgtg 46320
ggacgaagtc aggtttcggc gcaccgcgag cggatcgtgg ggcgcacact ggagttcctg 46380
ctgacccggg gggcgggcgt cggatcgagg aggcgacgtg ctggcaggat gtgtcgagtg 46440
gcccgcgcgc gtcgcggcgc agtaccgcgc gaaggggtac tggcgcgatg agaccctcgc 46500
cggcctgttg cgcagatggg cgcagagta cgggcagcgc acggcagtgg tgcgtggcga 46560
ggtgcgttg agctacgtcg aactcgaccg gcgggcccgc gccttgccgc ccggtttcgc 46620
ccagcggggg atcggttcag gtgatcgagt cgttgtgcac ctgccaaaca tccccgagtt 46680
cgtcgcggtg tgcttcgcct tgttcgcct cggtcggtg ccggtgttcg cactgccacc 46740
gcaccggagc gcggagatca ggtacctgtg cgagtactcg gcggcgattg cctacgtggt 46800
gcccgaatcg caccgcaaaa ccgactaccg gctgatcgcg gcgggcgtcg tctcgaccgt 46860
gccatcgctg cggcacgtgt tcgtggtggg cgatccgggc gaattcaccg cgtcgcgcga 46920
tgtcgacgct gaaccgcgc ccgtgccga acccggtccg gcggacgtgg ccttcttcct 46980
gctctccggc ggcaccaccg gactgcgcga gcttattccc cgcaccacc gcgattacgc 47040
ctacaacgtc cggatcaccg cggacaacgc cggctctgat gtcgacacca cctatctggc 47100
ggtgctgccc gcagcgcaca actacgcgct gggctgcccc ggtgtgctgg gtgcgctgtc 47160
ggtcggcggc acggtggtgc tcgccgacag cggcaacccc gatgaggcgt tcaccctgat 47220
cgaacgggaa aaggtgacga tgtccgcact cgtcccgccg ctggcgcgtc tgtggaccga 47280
caccgcgacc gagcgggatc tgtcgagcct gcgcctggtg cagggtggcg gttcgcggtt 47340
cgaggccgca ccggtgcga ggttcggggc tcggttcggg tgcgtgttg agcagtcctt 47400
cggcatggcg gaagggctgc tgagccagac cgcgctgtcg gattcagctg aggtggtcga 47460
ttccacccaa gggctccgc tctcgccga tgacgaggtc cggctggtcg acgccgacgc 47520
cgacgcgcgc gccgacgcgc accgcgacgc cgacgcgcgc gacgtcgcgc tgggcgaggt 47580
gggccacctg ctggtcgggg gtccgtacac gatccgcggc tactaccacg ccccggaaca 47640
gaaccagcgg gacttcacgc cggacgggtt cttccgcacc ggagacctag cgcgcgcac 47700
tccggaaggg cggctgatca tcgagggccg cgtcaaggac gtgattaacc ggggtggtga 47760
gaaggtggcc gccgaagaag tcgaaggtca cctgctggcg cacgaggcgc tgcgggaggt 47820
ggcgtggtg gccgtgccc acgcggtcat gggcgagcgc acctgtgcct tcgtggtggg 47880
ccgggcccga gtcggcgcgc ccgagctcaa ggagttcctc cgggggcgcg gcctggccgc 47940
ctacaagctg ccggaccggt gggagcaggt ggagtcgctg cccgaacgc ccgtcgggaa 48000
gatcgacaag gctctgctgc gcgagcaggt agcggcgaag gtcagccggg cgtcggccag 48060

ES 2 661 447 T3

cagcaggcgt cgccgggcca tccgatgagt cccgtggtcg tggtaggtgc cggcccggtc 48120
 ggcctgacgg ccgccctgtg cttgtgcgca cttggcatcg aggtcacctg gctggagaag 48180
 cggggccgcg atgcggtccg ccccggcagc cgggcgctct tcgtgcacaa cgattcgctg 48240
 cagctgctcg accgggcccc gcgcgggctc ggcaccgaga tcgccgagca cggcaccacc 48300
 tggcacaccc gccgcacgct gtttcgcggc cgggaggtct actcgaagac ctatcggaag 48360
 ccgaccgggg ccgaactgcc gccattcgcc agccttcgcc agatcgacac cgagcgattc 48420
 ctgttgctg cgccggagtc ggcaggtgct gacgtcgtct gggacgccga tgtcacccgg 48480
 gtctcggcaa cggacgccga cgtgacgttg gaggcggggg acggggcgag ttggcgagtc 48540
 cactacgtga tcgccgccga cggcgcccgt tctgcggtgc gatcgccggt cggcatcgcg 48600
 ctggcgggct cccgcgcgct gggttccac gtcgtggtcg atgtttccgc cggcgcgcg 48660
 ccccgccgcg tggagcgag ctatcactac gcgcaccccg cactgtccgg gcgcaacgct 48720
 ttgacgggtc cgttcacggg cggcttcag gtggatctgc agtgccgacc tgatgacgat 48780
 ccggacgacc tgaccgcaga ggtgcggtc gcgcgctgga tcgcgcgctg cgtcgacgag 48840
 gacgagatcg gcgacgtgct ctgggtctcc cactaccagt tcatgcaggt gatcgccgag 48900
 cggttcgtcg acgcgagccg ccgggtcctg ctgcgcgggg aagcggcgca cctgttcccg 48960
 ccgttcgggg ccccgggcat gaacagcggg atggccgacg cgggaagcggc ggcgcgtgcg 49020
 gtagggctcg cgctgacctg cggcaagcgc acgtctgtgg aggactacga tcgggtccgg 49080
 cgagaagctg cccgcgcgaa cagctgggcg gtcggtgcag cgttgccgca cgtggcaccg 49140
 ccatccccgt tcggcaagat gaagctggcg ctggccgcat ggctggctcc gcacgtcccc 49200
 aagatcgggg agtggtcga acgggcaccc tacgggcccc ggccggcccc gcgaccaggc 49260
 ggaagaggac ggtactgacc atcgcgacga aacgcgggga ccgagtgctc gtgtggcaca 49320
 acggtgctct ccacgaggag tccgctgaac ccaatcgtcc actgtgggtg gtggattcgt 49380
 ggctgctcga cgagggcgag gtcggtggcc tggcgctgca ccgagatcgg ttcgtcaacg 49440
 gctgcccgca agcggacaag aaggccccgg acgaggtgct tggcttctgg ctggagatgt 49500
 gccgtgcgtt gccgaggacg ggcaggtggt ttccgcgcgc ggagctcagc ggccatccca 49560
 agaagctccg cgcggtgata cgccttgcc ccgacaccgg gggagaagtg gtgctgtggc 49620
 cgtgcctgga cccgcgcaac tcgcctcgcc gcaaggggccc ggacgtcccg gcgctggcga 49680
 acgcgcggca gaacgcgctc gtagaggtg ctcatgacgc ggtgctgcta acgaaatccg 49740
 gagcgctgct tgaggccgct caggtgacct tgctgtggtg gagcgcgaa cgcattctga 49800
 tgccggtgca tgacgacttg ctgcttccca gcgtcaccct gcgcctgata cgtcaactcg 49860
 cgagcccgca cggcgtccca atagcgcgcg ttgacgcccc gctgcaggac ctgcggaact 49920
 gcgaggtgtg gtatgcgaac gccctgcacg gcattcgccct agtgcgctcc tggctgggaa 49980
 gtccagtcgg gctcggggaa tccgtgcgag cggcgcgctg gcaggacgcg ctccaccgagc 50040

ES 2 661 447 T3

tgcgtgagcc gctgccctga ggggcgcggt gcacgcgcac gacgctgccc ttgccgtgaa 50100
 tccccgggac cgggtgcccc gcattcgacc agcgcgggga acttccccgtc ccgatcgttc 50160
 ggggtccggg tcgatccccg tctccttgtc ccggtcgctc attggcagcg gtcgagcaac 50220
 atgcgaccac aaccggccaa aaacgaaacc agcgggctca ccttccggca aaccgctgt 50280
 tcacagccgt cgagacgacg ggattgtctc gacggcgctc ttacgcgct gatcgacacg 50340
 accggatcaa cggggtgctg tgccaacgga tgcacggtga cggccatgta gggccacacg 50400
 ggcagggcgc tgatgatgtc gtggagctct gttgcgtcct cggcgcgcca gagtccccag 50460
 ttctccccgc gggccgggac tcgccacatc cggacgagat gtcccttctgc tggccgctcg 50520
 gcggccatct tccgctcgtc gatcgagagt tgccggatcg tgcggccga aagatcgctg 50580
 ggccactcga ttttgatgtt gaccagaaat tccttcatcg aagtctcaa ctcctcacca 50640
 tttgtttcac caatgtcgcg gcgttacgtc ctgggtgctcc ggtgacgtat ccgccgggcc 50700
 agctgcccgc gccgccacg acgattccca tcgaccgtgc ggcgcgcgtg aagtcggttc 50760
 cgggcctgcc gtccagcagt tgtgacgga gcatgtcgcc gtgggtgatg ttgcttccgg 50820
 tgagtccgaa ctcatcttcc aggtcggtg gcgagaacca ccgcgcgtcc tcgattcttg 50880
 tctgaaggcc ggggaaggac tcttccagct gtgccagaca ggcgccgaag aaagtcgcac 50940
 catgctcgcg ccagtactcc cggccgagtc ggtacggcgc gtgccacacg ttgatgtca 51000
 tcaggtgttt gcccttgggt gccaaacctt ccgacgtgag actcgggatg agtccccaca 51060
 tgatcggtt gggcgaaggg cggcctgcga tgccgtcctc gactgccgcg gtgatgtaag 51120
 ccgggtccgg ggcgacccg aactgggccg agagcagtg ctccagtggt acgtcctcgg 51180
 gagcgccggc gacgccgggg agcgaatcga gagccagggc tatcttgaag gccgaccct 51240
 cgggcgggtt gggggtcgcg gcctccgtga gcgccccgga tggaagcatg gtggcgaggc 51300
 gggacgggtt gactgcggac accaccatgt ccaccccggt gatctgttca cccgactcca 51360
 gcaccacggc cgtaacctgt ccgtgttcgg ccagctccag acggacgaca ggagcgtag 51420
 ttctgatctg gacgccgtgc tcgcctcgcg ctgcgcgag tgcgtcgacg atcgactca 51480
 tggagccgac cggtagcccg accgagccgc gcagcgcgcg atcgccgatg ccgagcacgt 51540
 cttcggtcga cgaggctcgg gagatgggtc gcatgagcag tcccatagcc gaaccggcg 51600
 catccggaga gacgagttgc ccattgaggg cgagcatcat catcaagctc ttggcggcgt 51660
 cggtcgcaa cgactcggcg agcaggtcgt tcaggcttcc gtcgagcacg acaaggcgga 51720
 actcgtcgcg taccgtgtc ggcatccggt ccagcgcttc ctgcaacgag ggcggcggtt 51780
 cccacaacga gagcccgctc gggcgccga gagcatcgag gcgctcgacg agcagacgat 51840
 gccgttggc ctcaccggta gcgaacgtct ccagctgtgc ggccaccga tcacggtcgc 51900
 gccagccgac gaacaggccc tcacccatgc gctgcaacag ggtcacgtcc gggctgacga 51960

ES 2 661 447 T3

accgcaaccc gtgtcgcgcc agctccaact cctggatgat gcgtccctcc aatgaccccg 52020
 gggagttggt gatcgcgccg aggtaccgag gaaggtactc cagccgaccg gccatcccg 52080
 cgagcgtggg gcgagcctca aggaggtga cgcgcatccc gaagcgcgcc aggtagtgtg 52140
 cagcgaccag gccgttgtgt ccccgccga cgacgaccac atggggactc atgcggacgg 52200
 gcggtcgaag gtctgctcga acagggcgag ccagttgcg ccgagaacca ggtccagatc 52260
 ggcggcggtg aaccgctcc gcatcatggt gctgacgagc gtcgggaagg cggcagggga 52320
 ctggaaccac gtcggccact cgaccaccgg ggcgctctcg tcgacgga cggggctttc 52380
 gcgacccac cgcccatcc gccaccatct gacggcctcg acgtcatacc cgggtgtagaa 52440
 gtcggtgccc agggcgacgt gttccacacc gaaccgctcg acgggtccacg aaatcatgtc 52500
 cacgaattcc tgctcggtgg tcttgacgcc gttcttgagc atccgcacgt acgggctcag 52560
 gccacaacg ccgccggcg ctcgagagctc gcgcagcacg gtggacgact tcgggcgccg 52620
 gccgagttcg acatcggtac cgacgaactc cagcgggttg gcatgcgtga tcgcaaccgg 52680
 aacctggctg aactcgatgg tgtcgaggca cgtccgctcg gagcagtgcg acaggtcgac 52740
 gagcattccg acctggttca gttccttcac gaactgccgt cccacgtgag tggacaagcc 52800
 ggcgtcgtcg tcctccagc aaccggcggc gatgctgttc tgggtgttgt aggtgagttg 52860
 cacgatccg acgcccaggt cgtgaaagac ctgcaccagc cggatgtcgt cttcgaaggg 52920
 ggaggtgttc tggaatccca gaacgaccgc ggtacgcccg gaggagacca cctcacggat 52980
 ctcgcccgcg gtgcgggccc gcgcgatcag gtcggcgctc tcggcgaaga ggtcgttcca 53040
 gtggcccagg acagacaatg tttcccgggc ggtctccac accgcgaccg ttacgtggac 53100
 gcagccgacg tgcccgctgc gccactccaa gaaccgcttg cgggaaggcc tgctgtactc 53160
 gagacaatcc acgatgatgg ccccggtttg cgctcctgct gtcattgtcct ctgctttctt 53220
 tccgagtcgg attcttgttg aactgattt gggtcaggcg agtccgaaag ggtctcgggc 53280
 tgaggcaaa tcctgcacaa tcgctttgac ctcttggaac tcgcgcaagc cttcgatacc 53340
 gttctcccg ccatgcccg attgcttcac acctccgaac ggtgtctgcc agtccatgcc 53400
 ccggtaggtg ttgatccaca cggttcacgc gttcaggctc tgcgacatcc ggcgggctcg 53460
 cgccatatcg gaggtccagc agccggcagc gagcccgaag gcgagtcgt tggctatogc 53520
 gactgcctcc tcctcgccg cgaactccac gagcgtcagc accgggcca agacctcgg 53580
 ctggcagatg ccggcgctgt tgtcagctac ccgaaggacg gtcggcgcca cccagtaacc 53640
 ccgctcgagg ccggcagcca ccggtggggg ctcggcgccg cccgcgagca cgggtggcacc 53700
 gtcggcgatc gcctgggaga tcaacgactg gatgcgctgc tgctgtgga ccgaggcgac 53760
 cgggcccaga tgtaccccg cctccgtggg ggcaccgacg cgcagccggg acacgcgctc 53820
 cagcagcgca gcggtgaatt cgtccgccac accgtgttgc acgagcacc gagagccggc 53880
 cacgcaggtc tgcccggtcg cggaagacc acctgcacg accgcggtca ccgcatcgctc 53940

ES 2 661 447 T3

gagcggcgca tcatcgaaga tgatgtgagc cgacttgccg ccggcctctg cgacgaagtg 54000
 gccgagatga ctcgagacag acatcccagc ctgcgcagcc gcggctgtgc tgccggtgaa 54060
 cgagaccatg tccaccgact tgtgtcaaac gagccgctgc ccgacgccgg tgccaccctg 54120
 gacgatggtc acagcaccgt tgggtacgcc ggccctcggtc acgatctgct tgaggaggat 54180
 gctcgagccc ggcgagttaa ggcggggtt gataatgacg gcatttccgg aagccaacgc 54240
 cggggcgagt ttccaggtgc cgaggagaa tgctccgttg aatggcgtaa taagtgtcac 54300
 gactcccacg ggctcatacc gcagggtcgc ctcccttggtc gccgagatgc tgaggcggtc 54360
 ctcacgcgta gtgcgcagca gcgcggcgta gtactgccac cagcgtgcgc tgacggctac 54420
 ctccgacctta gtgcgaggaa atggcttgcc attggccagc gtctcgagcg ccgcgatttc 54480
 atcgacgcgg gcgtcgatgc cgcgcccgat cgcacgcagc acgtcggcgc gctgaagggt 54540
 gctgaccgag cgccattctc cgctggcgta ggtcgcccgg gcaacagcga cggcctcgtc 54600
 gaccagatcg acgttgtggc tcaccgtacc ccagggtctg ccggtcgtctg ggtcagttag 54660
 tgcttcgtgc cgggtacctt gacgtgtgcg gccgttgatg tccgagtggg attcggttgt 54720
 ggagagcaga gcgaataggt ctggcgaggt cgttgcaatc atacagactc ctgtgcgggt 54780
 ctcgatatga attcagattt ccgagccgga aaaggcgca atgaagggt caaccgggtc 54840
 gggcacgcgc tggaatcgtc ctgcggatta cctatcggct gctgtcgacc gcgcaccctt 54900
 ggtcggcact gatgcgcgag ctgcacgtgt ctccaccgtt cgtccgcctt gcggaatttc 54960
 atggccgggg acggagccga gcagctttct cgtgtagggg tgctgcgggt tttcgaagat 55020
 ctcggttgtt ctgttggctt cgacgacctc gccgtccttc atcacacaga cgtagtccgc 55080
 gatcaaccgc accactgcga gatcgtggga gatgaacagc aagctgaggg cgagctcggc 55140
 ctgcaaccgg tcgagaagcc ggaggatctg gcctgcacc atgacgtcga gggcgacac 55200
 gggttcgtcg cagatgagca gatcagggtt gagggcgatc gccctggcga tcgcgacacg 55260
 ctgcgcgtgt ccgcgggaca gttccggggg gcggcggtcg atcacgtccg aggaaagtgc 55320
 gacctgatcg agcagctcga gcaccgcgc gcgtcgttgc gccttgctc cgcgcttgta 55380
 gaaggccaat ggctcctcaa cgatcctgcc gatcgtgtac atcggttca gcgaggcgta 55440
 cggatcctgg aataccggct gcacacgctg ccggaatccg gcgagctggg ccttatccat 55500
 cgtcgcgac tcttgcccct cgaagaagat cgagcccga gacggctcga tgagccgaag 55560
 gagcatccgc gatgtcgtcg tctttccga acctgactcg cccacgacgg cgaccgtctt 55620
 tccgccggg atgtcgaacg tgacgcgggc agcggctcgg acctccttgc gattgcctcg 55680
 gggacggtag attttggctg cttcccggac ctccacgagg ctgggcttgc cagcgtccgt 55740
 ctcaccgggtg gcagtatggc ccgagtcgac gcgcacggcc gcgatgctcg gggcggtcgc 55800
 gagcagggcc ttcgtgtact ggtggcgcgg agttgccaga acctgcgcag tgggccctgc 55860

ES 2 661 447 T3

ctcgacgata tccccctcgt tcatcacgat gatccgctgg gcacgttcag ccgcgagggc 55920
 gaggtcgtgg gtgatgagca ggaccgcggt gccgatctcg gcggtcatcc tgtcgatctg 55980
 gtcgagaatc gtccgctgga cggtgacatc gagcgcaactg gtgggctcgt ctgcgatgag 56040
 cagcttcggc cggcaggcga gcccgatcgc gaccagcacg cgttggcgca gcccgccgga 56100
 cagctcgtgc ggatactgcg cggcacgctg ttccggctcc gggattcccg cggctcgaag 56160
 tgcttcgtgg ctttttcccc ggacgtcgcg cctggctcgc aaccatgtg tcagcagggg 56220
 ttccgcgacc tgggtacctt tcttcatgct cgggttcagg ttccgacatc ggtcctgcgg 56280
 caccaaaaccg atgagacggc ctcgtagcgt acgaagtctg gtgtccgact tcgtggtgag 56340
 gtcggtgttc tcgaagcggg tgctaccggc gcggattttg ccgttggcgg cgagcagccg 56400
 gatgacggac atggccattg tggacttgcc ggaaccgat tctccgacga tcgcgactgt 56460
 ctgcgccggt ccgacacaga acgacacgcc cttgacaacg tcggaggaac cgttggtgcc 56520
 ggtgaaactg acgacaaggt tgtcgacagc aagcagcggc agcccgctcg gggagtcact 56580
 catgggtttg ccccttctcg ttggtgacgc tcgccgggct tttcccaatc cgcctcccg 56640
 gcgctgggcg gcggtgtga cgaactcccg atgacgagct tcggggggac ggtgatcggg 56700
 ggcacgcgat tgccttgtgc ctgcgcccta accaagtga tgagcctgtt ggtcgccgct 56760
 ccggcgatct cttggacaga gccgctgacc gaggacagcg gtggcacgag gtacgcgcg 56820
 aagtcgaggt cggcgaagcc gacaacggcg atatcagtgc caacggtacc gccgaggtcc 56880
 gtgatggctc ttatcgcgcc caccgaggtg tggctattcg acgcgagcac ccccttcggc 56940
 caccgaccga cgcgcaggtg cgcgcacatg gccgagtaac cggccttgta gtcgtagccg 57000
 tcggtgacga tacggatccc atccgcctgc ccgtggcggc ggaccacgct gctaaatccc 57060
 tgccaccgat tctcagcggg gaccgagccc ctgatcgcac caaggaatac aagcggcaca 57120
 cccagtcct cgcgaggtg ggcggtggcc tcttcggcac cgcagtagtc gtcatagaga 57180
 atcgtcggga cgtccaggcc cggaatggtc tgtccgatcg tcacgacatg aatccccgc 57240
 tcggcgatcc tcgccagctg tgcgcgatc tcgggactgc cgaggtcgtc acctgtgttg 57300
 atgatgattc cgtcgacgcc ctcgtagcgg atgcccgcga ggtagctgag cgccttggcc 57360
 tgatcgtggc cgaagctctt cagctgtagc gcgaagctgt gctctgcgca caccgtgtcg 57420
 accgacttcg cgatcatcga ataccacggc tgcgagatgt caccgaccag gagggccaga 57480
 caccggtggg agcctgatcg caggccggac gccgccgat tggggtggaa cgacagctcg 57540
 tccgcgacct tcagcacctt cttccgctg ttctcgcca ccagctccgg acggttgagc 57600
 gcccgggaca cggtcgagat ggaaaaaccg gtggcttcgg cgatctcgcc catcgtgccg 57660
 gcactctggc tcgtcacagc ttgtctcgcc tcccggcctg ggccgctagt catgcgatt 57720
 ccccgagaac gtcaccagca cgaacatctg gtctccctc aactgagctt gatcttcgga 57780
 ttgaagtagg cgtacaggat gtcaaccagg aagttgatca atacgaagat caccgcatc 57840

ES 2 661 447 T3

aacaggatgc cgccctgcaa cagcacggtg tcgcgctgag tgatcgctg ggacaaggcc 57900
 agacctatac cgggccatgc gaataccgtc tcggtgagca cggcaccgct gagcaggaag 57960
 ccagcctgga ggccgacgac ggtgatcacc gggggcatcg cgttgggcaa cgcagccga 58020
 aacacgacct tgcgctccgg cgtaccgagg gatcggggcg tcgacacgaa gtcgctgctc 58080
 atgacatcga gcatggagcg tcgctcatt ctgcgcatga ccgctaccga ccacgcccgc 58140
 agcgtgagca caggaggat cgcgtggctg gcgacgtccc ccagtccacc cccgcccagc 58200
 atgcttcgca taccgctaac cggcagccag ccgagggtca gaccgaagag caagacgagg 58260
 atcagcccca gccagaacga cggcatgctg gtcagaacaa gcacgaacac cgtcaggccg 58320
 cggtcgcgca acgagtggc cttcagcgcc gccagctgc cgagcacgat cccgagaacc 58380
 gccgccacga ttacggcggg aaccgcgatg acggcgggtg tcaccagccg tgacaacagc 58440
 acgtccagag caggctgatg ataggccgtc gagacgccga gatctcctcg aaggagatgt 58500
 cctaggtaga tcaaatactg gacgaagatg tttcgatcca cgccgagctg gtcacgcaga 58560
 gctgtcttct ccgcctcgct cgcgttgccg cctagccgca catcggctat atcaccgga 58620
 atgatctggc tgagcaggaa gaccgcgaac gtcacgccga agatgaccgg gagcagctgg 58680
 agcgcacgtt tgcctgtgta cctgagcatg ggtgtgcctc ctccaggagt gttccgcact 58740
 cgcggatcca ggtgtgcacg aacggcgtcg ccgaccaggt tgaagccgac agtgagggtg 58800
 acgatcgcca gcccgccgag cgtcgccacg tggggcgcgc tcaacaaggc tgcccgccg 58860
 ccgttgagca tcgctcccca gtccggcgtc gggggtttta cgccaagccc caggaaactg 58920
 aggcctcgcc ccaggatcac gttcttgccg gtctgcagcg cgccgatgac caccagcggg 58980
 ctcaaacgt tgggggcgac gtgcagggtg gccgtccgga agtgtcccag ccccagcgtg 59040
 agagccgact cgacgtaggg ctgctccttc tcgcctagga cgcttcgag cagcagacgg 59100
 ccaacctcgg gaatggtcac cacgaccatc gcgatgatca tcgagaccat gcccgcccc 59160
 agggcggcgc cgatccccag cgcgagcagg acgcccggga atgogaagac gacgtcgaac 59220
 acgcgcataa tcaccgcgtc gatcttgccg ccgacgtacc ccgacatcag gccgatcagc 59280
 ccgccgatga ccagcgcgat cagtgtcggc gcgaccgcga cgaagagggc gaacctggcg 59340
 ccatgaatca gccggtcaa cacgtcccgt ccgagttcat ccgtcccag gacgtgtccc 59400
 ggggagaaca gggccagata cctctggtgc gggtcggcca ggatcgggtc gtgcggcgag 59460
 ataaggtcgg cgaagatcgc acagaacacg accacaccga cgatgaccac gcccgtaga 59520
 tgggtgggct ttgtcaatgt cgaaagccac ccgctgagcg cgggcccgtt tgagagcgcg 59580
 gtttcctgga caacaccagc cgccatctcg atcatcctct cggctcttga cagttgaact 59640
 caggccacag tccaaagtgc gccgagggtg gggatgacgt ccggcaagcc gagcgtgcgc 59700
 aatgcgtgcc agagcgtcga ctggttactg ctgaccactg cggcaccggt gtccatttcg 59760

ES 2 661 447 T3

aggtagggca atgtctcgaa ggttcggtag gtgccacagg agatgaggat cgcgtcagca 59820
 ccgccggcct tcttgtagac gcccttgccg aggcggtagg cactgtcggc gtcgtgggtg 59880
 ctgcgctcgt acgggtcggg catctggagg ccttcgatgg cactcacctt gatgccgtga 59940
 ctttcgagag acgctgccag cgcctcgttc acatcggttg tgtaaagtgt tccgacagcg 60000
 acggtgggtg cgtaagcgc ctgcagcgcc ttgaccgcg actcctgcat cgtaatggct 60060
 ggaatcccgg tgtgcgccgt gatctcgtg gtgatgcgct cgtcggcgcc tgggccatcg 60120
 aggaaggtcc ccggcgcgcc gcatcgacg atgaggtcga ccttgagggt cgccagctgg 60180
 cgggacgcct ccagcacctg cgacatcatc accttcagac cccctcgtc caccctgggg 60240
 acgatgggtg gggcatcgcc gaaggcgacg ccggtcgggg cgacgctccg cactcctcg 60300
 gcgccctcga tagcgggttc gggcggtatc tggccgatct tggcgcgcca gcctagaagc 60360
 attctgaacc caactcctca agtcgtgatt catcggcaga ccctgtggcc cactctggtc 60420
 gcttccccga taccggtcag aacgcgaccc agagggattt gagaaccgtg gcgcagagac 60480
 tgacggggga gtgcgacgac gaggaccata agcgcactgc acggccaatg tcaacgtttg 60540
 cagtcacgcg gtcaaaggca catgcgcccc gcttgagccg gcgacggaca cgtcgagggg 60600
 cggccccgct ggaagtcgca cacgtccgc gactgatcga ggccaggtgt cgcagcgtct 60660
 cgcgagaacc gttgtgcgac agcggaaatg gagcttctac gtagcccaa cgggacgggtg 60720
 tgcacgggtc tgctgaggcg agcgcgcgct ttgatgggtc cgccctcggg ccgttgacag 60780
 cctgaggcta gccatctacg ttatccccctc acgccacaat cgtttgcaa gaccacggag 60840
 tcccgcacga caatgcgcaa gaggcgtgtc ttggccgccc ttcttgctgc cacatccctt 60900
 ctgcgctcgc ccgcctgctc cgtgtgtgga cagggcggcg gcgaccaacg ggttctcgtc 60960
 atgcctcggg gaggggagat tccggcactc gaccgcgaga gcatcagtgg aaccgtcgcc 61020
 cttcgtgtca ctgacgccat ctacgagact ctcgttcggg aggatctcga tacgcagacc 61080
 aagacggctc ccgaagtcag gcctgccctg gccgaatcct ggaacgtgtc ggcggacggc 61140
 cgggtttaca cgttcaagat ccgggacggc gtgacgtttc acgacggcac gtccctggac 61200
 gcagcggcga tcgctacgaa cttcgaccgc ctgatggacg accgctctcc ggtgtacaac 61260
 cagatcgcga gcgccaacat gaagtctgtg acccgctgga tcgcgagcgc caaggcgacc 61320
 ggcgatctta ccttcacat cagcgtcaag aagccgttcg cggagcttcc gcggctgctc 61380
 accgaccgcc ggatggcaat catcagcca gtggccctgg cgaagtaccc aggtgatgca 61440
 gtcggtcgac atccggacgg caccggaccc ttccgaatcg agagtgtcga gcagggcacg 61500
 gatctgcac tcgaggccaa tgtgacgtac tggcgcgcg ccccgaaagt cgatggcctc 61560
 gtattcgtga acatgcagga tcccacaacc acggcgatgg cgctgcagac cggagaaatc 61620
 gatgtgattc ccagcgccag cgcggagcag atcgcgcagc tgcaaggcaa atctgcagcg 61680
 gtggtgcagt acccgggccc gcccaaccag tattttatcc ggctcaaac caagacagcg 61740

ES 2 661 447 T3

ccgactgaca acccgagtt ccgccaagcc ttgaactacg cgatcgaccg cgacggaatc 61800
 atcgctgccg tggatggcat ggggcagatc atgggcggtc cgatcccgtt cggcaacgac 61860
 gccttccaag aaggcgctcg gcagcgatac gagttcaatc cggacgaagc acgccgtctc 61920
 atccacgctt ccgggggtcca gcttccggcg gtcgtaagg tatacgcccc gagcggcggt 61980
 cctggcttct cccaagcaag ccagatcatg tcgctcgtc aacaagatct gaagaacgtc 62040
 gatgtccagc tcgacgtcga atacgtcgag ttacttctgc tggtaaacct ggaggcgctt 62100
 gggtaacaaga acggcgctcaa cggctccttc aatgggtgga ccaccggcgc ggacagcggg 62160
 tactggctgg agcgaatgtt tgggtcggac cagataccgc cggccggtgt gaaccgcggc 62220
 tggtaactga acacagatct agctcggatc ttccgaccag cgcgtaacag catcgagccg 62280
 gagacgcgcc gcggtctcta tcgcgaagcc gccgacatca tcgctgacga cgcaccttgg 62340
 gtctttctct accaagaccg cctgcgcgg atcttccgg cgaacatcac gggaatccac 62400
 gggacaccca gcgtgttcgt cgactacgcc actatcagca agtcgtgacc aaacgctgag 62460
 caaggcacac agcgtttcgt ggcggtggac ggaatcgctg tcacgaatcc gtgcggtgtc 62520
 gcaggctccc tgccaaccgt tgaacggctt cgccatgtcg ggttgcgat gagatcgctt 62580
 tgtgaatggc gtacggccca tacggatata cagcagcagg atggcctgca ggccgagatg 62640
 agcgggcatc tcggctatgc caagggtgat ccggcagggc ggggcagccc gaactccgc 62700
 aggcggcatc agcgccaaga agcatcgccg ctcccgggca ttccggacgc ggtgatctt 62760
 tacctcgcc aaccagctcg ggttgatccg tcgtgatgga accgtcctcg cggccagacc 62820
 gtttagggcc tggcggtagg tgccaaattt cggcggtgaa cggcgctcgtt aagccctgtg 62880
 cgggagaacc gcatgcacgg attgacgggc gggagctgga aacggaacca cgatgagacc 62940
 acggtcatgg agatcaacga tccaacggga aaccgttggg tcaccgatgg ctccgcgacc 63000
 taccgccgat caggttcacc gcgccagctc ccgaccctac tcggttttcg caggtcgggtg 63060
 gcctcatagg ggtttactgg ggtcaagggt cgcaggttca aatcctgtcg gcccgatgtt 63120
 gccttttcgt gtccctatgt ggatcttgcg gtgttcgcgc gttcgagatc aaatgggcac 63180
 tggccgggtg cgcggtgcgg gccgcgccac cgcgggtccc ggtcgtcagc ccgccgtgaa 63240
 attcggcagg tcggcggtac gggccctgtg caggttcttg atgacgttac gagcgtttga 63300
 gtacttggtc gccccggact cgtagagcag ccatgtcgtg ccgttgtagt cggcgatgcc 63360
 ctccgacatg ctcgcgccct ggaagcactt ggcggagcgg tcgttgatgt tcctgccgtg 63420
 ggcgaccacg tagatgttgc tggggttctt ccggccgtgt gacgtgctgt agacgaagtg 63480
 cgtgtccgtg accatcacgc cctgcgtctt gtccgggtgt acgtaggctg cagtctgctt 63540
 gagcttaccg tccttgttca cctttagtag gtacatccgg tccttttcga aatggcccgt 63600
 gtacagctcg tcgccgtagc cgggtaggaa cgacgctcct ttgacctctt gggccggacc 63660

ES 2 661 447 T3

gttcggcgcg agcgactgcc tgcccggggc cttgaacttc gccgcgaggt ccgtgggtctt 63720
gtacttggtc atggacctgt tggtgccacc cggtacgaac acccagccat ttacgaccgc 63780
gatgcccccg gcgtgcttgc ttttgatgtc gaccgggtccg aggactctgc tggcgccgc 63840
ctcgatcccc tagatgcgcc ccgtgtaatc aggaaccttg ggatgatctg tgtacatgct 63900
gatcagcatc aggtcgttgc cggcgccgtc ccagttctcc caaaccccca ggccctgggg 63960
gatgtaacct tcggacaggc cgggtatcag gttggcctgc ttgaacttgc tgtcgtatac 64020
cgttccggcg gaagggccgt ggtagggggc ttttcctccc gcgcgctgc tgttgacgt 64080
gatcgcgcg gcttcctgcg cgaccgcggt ggccgggg 64118

<210> 2

<211> 6942

<212> ADN

5 <213> Saccharopolyspora Spinosa

<400> 2

atggcagatc agggatatcc gcgatctcgt gccgaagcgg cggcggtgtt cgaagcggtc 60
gaggacagcg tcgcggccct gacgcgttcg accacgtgtg ccgcgtggac gcgggaccgg 120
gatgatctgt tgctgggtgat cgaggagccc gaaggtggtg ctccagggga cttcggcgcg 180
ctcgctcaag aggtgcgcgc ggccgtgtcg cggcagcacg acctcgtgcc tgcggagatc 240
gtgttcgtcc cagctcgatc cattcctcgc ggcggaacg acgaggtgct gcgcgcggag 300
acggctgcca ttcatctccg ggggtggctt gacgtgctgc acgcgaccg gcagctggga 360
gtcgcggcgc cgccgctgct cggcgcgga gaatccattg tggacgacga ggtgctgcgc 420
ggcctggctg acgcggccga aagcgtggc gaggcgcgc cgctggacct cgaactcatg 480
cggctggacg tcgccgagtt gctcggtgaa tcagcggaga gcatcgccga cgacgaagac 540
ctgatcatgc ggggcttgga ctccatcatg atcatgacct tggtaacca gtggcgcgcc 600
gacggggtgc aggtcagctt cccggagctg ttccaggacg cggttctggc caagtgggtg 660
gagagcatca gcgggaccag cgagcccag gccgcgacg aggcggcggt cgacgagggc 720
gtggtcgacg aggcggcgcc gttcgacctg tcggtgatgc agcatgctta ctgggtcgga 780
cgcgcgacg agcaggtgct cggcggggtc ggcgcgact tctacaacga gttcagcgg 840
cgcgacgtgg atccgcccgc gctggagcgc gcggtgcggg cgctgaccga acgccacggc 900
atgctgcgag cccggttcac cgatgacggc aggcaatga tcgccgccac ctgcccttgg 960
cccgtctgt cggtgcacga tctgcgtcag ctcccgcgg acgaggtcga ccaccgctg 1020
agctcgctgc gcgagcgcc gtcccaccgc cggctggagg tcgagcgggg ggaggtgttc 1080
gacgtccagc tctccctgtt gcctggcgg gccagccggg tgcacgtcaa catcgacatg 1140
ctggtcgccg acgcgttag ctttcgcatc ctgctcaccg acctggccag gtgctacgag 1200
gaaccggagg ccgagcagcc gcccatcggc tacagctacc cgaagtatct ggcgcagcgc 1260

ES 2 661 447 T3

aaccgccggc ggaccaggc ggccgagcag gcgcgcgcac actggcggga ccggatcgcc 1320
gacctgccag gcgggccgga gctgccgttg gccatcgac ccgaacggat cagccagcgc 1380
cgggtgacca ggcgttacca ctggctcgac gcgcaggggc gggtctgct ggccgccaag 1440
gcgcggcagc acgggctgac catgccctg gtgttcacga ccgccttcgc cgaagtgtc 1500
gggacctgga gccccagcaa gaggttcctg ctcaacctgc ccctgttcga ccgggaggcg 1560
gtgcaccccg acgtgttcctg cctcagcggc gatttcacca acttgatcct gctggaagtg 1620
gacctgaccg agccgagctc ctctcgctgag cgcgcggcgc ggctgcgtgc gcagctgcag 1680
tccgacacgg ccaacgccga gtactccggc gtgcgggtcc tgcgcgacct ggccagggcc 1740
aacggcggag aaccgggtgc cgcaccagtg gttttcacca gcgggctgag cctcggcgag 1800
ctgttcagcg aggaggtccg ccgctgcttc ggcagcccg catggatgat ctgcagggc 1860
ccccagggtg ggctggacca ccaggtcacc gagcacgacg ggggcttgct gctgaactgg 1920
gacgcggaga acgagctgtt ccccgacggt ctctctgacg ccatgttcgc cgcctaccgc 1980
gtcgtgctgg actggctggc cgcagccgac tgggcgcagc cggtgccgcc gttgctgccg 2040
gccgagcaag cgcgggtccg ggctaccgtc aatgagaccg aggggcccgc gtctggccta 2100
ctgctgcacg aaggtttctt ccgcactgcc gagatcgcgc cgcaacggcc cgcgttgctt 2160
tggggcgagc agggcgcttct cagctacgga gcgctcgccg accgcgcccg ccgcattgcc 2220
ggcttcctgc tggggcgcgg ggtgcggccc ggtgagcggg tggcgatcac gctgccgaag 2280
ggtcccgacc aggtggccgc ggcgctcggg gtactactgg ccggggcgac ctatgtgccg 2340
gtgggggtgg accagccggt ggcgcgccgg gcgaagatct accgcgcgcg cgaggttcgc 2400
cagggtactca cactgccga ggagcggacg gactcgacct ggcccgcga tgtcgaaccg 2460
gtcgcggctg aggtcgccac gtcggtggac ccgctgacca ccagggtgcc aggggaccga 2520
accgacctcg ccgcctacgt gatcttcacg tccggttcga ccggcgagcc gaagggtgtg 2580
atggtcagcc atcgggcggc gatcaacacc atcgccgacc tcaacgagcg ctctcaggtc 2640
ggcgcgaggg accgggtgct ggcgatctcc gcgctggact tcgacctgtc ggtctacgac 2700
atcttcggcc tgctctcagc tggcggggcg ctggtgctgg ttcccagggc cgagcggcgc 2760
aacgcgcgc actgggtgga gttggcacgg cgctggaagg tcacgatctg gcagtccgtg 2820
cccgcgttgc tggacatgct gttgaccgcg gccgcggctg cggaccaggc gctggacctg 2880
cggttggcgc tgctgggcgg tgactgggtg accgtcgacc tgctcgaccg cctgcgggcc 2940
tgctcgccga agggcaggct cgtcggtctc ggcgggacca cggaggcgc catccactcc 3000
accatttgcg aggtcgaaga gatcccggcc cagtggagct cggtgcccta cggcaccgcc 3060
ctcgcaacg tccgctgccg ggtggtcgat gagcggggcc aggactgcc ggactggggg 3120
gtgggggagc tgtggatcgg tggccgtggt gtgcgctgg gctactgcgg cgaccccgac 3180
cgcaccgccg aacggttcgt ggtgcgggac gggatccgct ggtaccgcac cggtgatctg 3240

ES 2 661 447 T3

gcccggtact ggccggacgg ttccgtggag ttccctgggtc gcgccgacca ccagggtcaag 3300
 gtgcggggcc atcgcatcga cctgggtgag gtogaagccg cgctcaacag ccaccccgac 3360
 gtgcgccgca gtgtcgtggg ggtgacgggt gcggccggta cccggctggt ggctgccgtg 3420
 tttcccggtg ggctgtgtcc gcctctggaa gacctgcggg actggctcgg caccagctt 3480
 cccgcgcaca tgggtgccgc cgacttcgtg gtgttcgacg agctaccact caacgccaac 3540
 ggcaagatcg accgcaagcg gctcgttcga ctgctggcgc cggaagagcg caccgacggcg 3600
 gagctgcggg cgccgcaggg tccggtggaa gagctcgtgg ccaggccctg gtcggatctg 3660
 ctccggcggg agccgatcgg acgcgacgac gactttctca cccgcggcgg tgacagcctg 3720
 ctggccaccc gcgtcgtggg caggctgcgc gcggccgggtc tgcgcggggc tgagctgcgg 3780
 gcgttggtcg ccaagccggg gctgcgggac ttgcgaacag acctcgaaat cggcgcgtca 3840
 ccggttgcca gcgcgtccc ggctgacctg gccaccggg acgagccgtt cccgctcacc 3900
 gacgtgcagc aggccttacct gctggggcgg ggccgggact tcgcgctcgg cgggtgcgcg 3960
 tcccactggt attgggagtt cgacaccgcc gatctggacc tgccgcgctt ggaggaggca 4020
 tggaaccggg tcgtcgcccg ccacgagatg ctgcgggtcg tgttcgagga cggccggcag 4080
 cgcacatctg ccgcggtgcc gcgctatgcc ctctccgtcg aagacgcgcg tggcgggtgct 4140
 gagcaggctg cgctggaccg aatgcgggcc gccatgtcgc accagggtgt cgacgtctcg 4200
 cgctggccgc tgttcgacat ccgcgcgctg cgctacagcg gcgggcgggt gcgcctgggg 4260
 ttcagcctcg acctgttgat gctggacgcg ctgagcatcg tgatcgtgtt cgcgagctg 4320
 agcgagctct accaccgtcc cggaaccggg ctgcccgaga tcggggtctc cttccgcgac 4380
 tacgtgctcg gtgcttgccg ggatcgagag cagctcgccg agagcaagcg ctactgggca 4440
 gagcgggctc cttcgtacc tccggcaccg caactcccgc tgcgcaccga tccggaactc 4500
 gtgcgcgagc cgcgcttcac ccgcaggcag agccggattc cgcggcaacg ctggcaacgc 4560
 attgtcgaac gagcgcggca gcgcgggctc accccggctt cgggtgctggc gacttgcttc 4620
 gcggaggctc tgggacgctg gagcaggcag cgggagctga cgctgaacct gactctgttc 4680
 gatcgcccg aggtccaccc cgagatcaat ggctgtgtcg gggatttcac cgcctgctg 4740
 ctggtgggtt acgagccgag ggccggggag tgcttcgccg agcgcgcggc ccggttcag 4800
 cgccaaatgt ggcaggacct cgaccaccgc gacgtctcgg cagtgcagg gtatgcgggac 4860
 ctggccgctt cccgcggatc tgcggccgct atgccggtg tgttcaccag cgcgctgggc 4920
 atcgcccggt acgtcgccgc cgagcggttc gccagggagg tttgggggct gtcacagacc 4980
 ccgcaggctt ggctcgatca ccagggtgcg gacgacgagg gcggcatgcg gttcaactgg 5040
 gacgcgggtg aggagtgtt cccggaagg ctgctcgacg ccatgttcca agctgagtgc 5100
 cgcctgctgg actggctggc cgatgcgcac tggctgcagc aattgccaga cctgcttccc 5160

ES 2 661 447 T3

acttcgcaac gagcgggtgcg cgagcaggag aactcgaccc gcgaaccgac cgccgcagcg 5220
acgctgctcc acagcggatt cttcgctcgt gcggacgccg acccggcgcg taccgccctg 5280
gtgtggggtg cggaccagcg gtgcgagtac ggcgaaactgg ccgacctcgc cctgcgggtg 5340
gccgcgcact tggtagagcg cggggtgcgg gcgggcgaca atgtggccgt tgcgttgccc 5400
aagggacgcg accagatcgt tgccgtgctc ggctgctgg ccgctggagc cgtgtacgtg 5460
ccggtggggc tcgatcagcc cgatcagcgc aggcagcgga tctactcgcg ggccggcggtg 5520
cgctggtgc tggacgacct ggcgagagcg taccgccgga agcccttggc ggggccgggtt 5580
gaggtggatc cggagagcct cgctgacgtg atcttcacct cgggttcgac cggtagagccg 5640
aaggggtgta tggtagacca cgctgcggcg atgaacacca tcgccgacgt caacgcccg 5700
ttcggagtcg gcgccgagga ccgggtgctg gcgatctccg cgctggactt cgatctgtcg 5760
gtgtacgaca ttttcggctt gttgtcggtc ggcggttcgg tcgtgctcat cgacgaggac 5820
ggaaggcggg aagcccgca atgggccggg ctttgccgc gctggaacgt cacggtttgg 5880
aacaccgtgc ccgctctgct ggacatgctg ctggtggtgg cagacgtcct gcccggtcg 5940
tttcggctgg ctctcgtctc cggcgactgg gtcggcctgg acctctgga ccgacttcgc 6000
gagcgggcac cggggtgcgc gctggtcgcg cttggcggtg ccaccgaggc cgcgatctgg 6060
tcgaacttct tcgaagtca acgggtcgac ccgcgatggc gttcgatccc ctacggcaga 6120
ccgctggcca accagcgggt ccgcgtggtc gacccaccg gtcaggactg tcctgagtgg 6180
accgagggg agctgtggat cggtagccgc ggtgtcgcgg ccggctactg cggcgacgcc 6240
gatcgaccg ccgagcgggt cgtcgaatgg gacgggatcc gctggtagc gaccggggac 6300
ctcgggcggt actggcctga tggcgtgctg gaggttcctc gccgggtcga caaccagggtg 6360
aaggtgcgag gtcaccgcat cgaactgggc gagatcgagg cggcggtgac cgcgcatccc 6420
gacgtcggcc aggcgctcgc cgcctccgtc ggggaccggg ctcccgcgct cgtggcggtg 6480
gtggtacccg agggcgaccg gtcctgcgat tccgacgaac tgcgatcgtg gctgctcgac 6540
cggctgccc cctacatggt gcccgagcgg atcgccgaga tcgctgaggt gccgttgacc 6600
gccaaaggca aactcgaccg cgcctcggcc atccgcgggc tggccgacgg cgcgcgggaa 6660
ccggctgaga ccggtgaacc gcttcgcggc gaggtggaga tcgctgctggc cgagcaatgg 6720
cgcgaagtgc tgggctgccc ggctcgcgc cgtgacgaca atttcttcgt actcggcggc 6780
gacagcctgt tggccaccgg cgtggtggag cgcattcgcc gcctgttcgg cgtggacatc 6840
gcgttcgcgc agttcctcgc cgggccatcg cttcagcagc agggggaatt ggtgaaacag 6900
cgttcgcagg ccggtcacca gatggaggaa ggcgtcatct ga 6942

<210> 3

<211> 412

<212> ADN

5 <213> Saccharopolyspora Spinosa

<400> 3

	gtctacctcc aggcgttcct ggagccggac ttcccgttcg aactcgccgg gatcgtcgcc	60
	cagggcagcg accggtccca tgcctgcgca gaggttccgt gtacaccaac	120
	gtcgaccagc tccccgacga catcgacatc gcctgcgtcg tcatcagcgc aggagtgcag	180
	ggtgggcgag gcgcggaaat ggcgcaggcc atcatgggccc gcggcatcca cgtgctgcag	240
	gagcacgcgt tgcaccacga tgagctggct gaatgcctgc ggtgcgcccg caaacacggg	300
	gtgcagtacc ggttgaacac ccactaccg cacatcgagc cggtcggccg gttcatcgcc	360
	gcagcacggg agctggctcag caagcagcgg ccgctgttca tcgacacgac ct	412
	<210> 4	
	<211> 20	
	<212> ADN	
5	<213> Artificial	
	<220>	
	<223> Cebador directo de PCR	
	<400> 4	
	gtctacctcc aggcgttcct 20	
10	<210> 5	
	<211> 20	
	<212> ADN	
	<213> Artificial	
	<220>	
15	<223> Cebador inverso de PCR	
	<400> 5	
	aggtcgtgtc gatgaacagc 20	
	<210> 6	
	<211> 303	
20	<212> ADN	
	<213> Artificial	
	<220>	
	<223> Secuencia de nucleótidos del sitio de unión a actinofago (attB) de Streptomyces lividans que contiene sitios BamHI y NdeI	
25	<400> 6	
	ggatcccgga atgtcgacga tgtaggtcac ggtctcgaag ccgcggtgcg ggtgccaggg	60
	cgtgcccttg ggtcccccgg gcgcgtactc cacctcacc atctggtcca tcatgatgaa	120
	cgggtcgagg tggcggtagt tgatcccggc gaacgcgcgg cgcaccggga agccctcgcc	180
	ctcgaaccg ctgggcgcgg tggtcacggt gacacggga cgtgcgacgg cgtcggcggg	240
	tgcggatacg cggggcagcg tcagcgggtt ctcgacggc acggcgggca tgtcgaccat	300
	atg	303
	<210> 7	
	<211> 22	
	<212> ADN	
30	<213> Artificial	
	<220>	
	<223> Cebador 2 directo de PCR	
	<400> 7	
	tgtaggctgg agctgcttcg aa 22	
35	<210> 8	

<211> 29
 <212> ADN
 <213> Artificial

 <220>
 5 <223> Cebador 2 inverso de PCR

 <400> 8
 catatggata ttccggggat ccgtcgacc 29

 <210> 9
 <211> 4477
 10 <212> ADN
 <213> Artificial

 <220>
 <223> secuencia de nucleótidos de plásmido pCR8 / GW / TOPO_aac (3) -attB

 <400> 9
 ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga 60
 taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcgggaaga 120
 gcgcccataa cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca 180
 cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaata cgcgtaccgc 240
 tagccaggaa gagttttag aaacgcaaaa aggccatccg tcaggatggc cttctgctta 300
 gtttgatgcc tggcagttta tggcgggct cctgcccgc accctccggg ccgttgcttc 360
 acaacgttca aatccgctcc cggcgggatt gtctactca ggagagcgtt caccgacaaa 420
 caacagataa aacgaaaggc ccagctctcc gactgagcct ttcgttttat ttgatgcctg 480
 gcagttccct actctcgcgt taacgctagc atggatgttt tcccagtcac gacgttgtaa 540
 aacgacggcc agtcttaagc tcggggccca aataatgatt ttattttgac tgatagtgc 600
 ctgttcgttg caacaaattg atgagcaatg cttttttata atgccaaactt tgtacaaaaa 660
 agcaggtccc gaattcgccc tttgtaggct ggagctgctt cgaagttcct atactttcta 720
 gagaatagga acttcggaat aggaacttat gagctcagcc aatcgactgg cgagcggcat 780
 cgcattcttc gcatcccgc tctggcggat gcaggaagat caacggatct cggcccagtt 840
 gaccagggc tgcgccaca atgtcgcggg agcggatcaa ccgagcaaag gcatgaccga 900
 ctggaccttc cttctgaagg ctcttctcct tgagccacct gtccgccaag gcaaagcgct 960
 15 cacagcagtg gtcattctcg agataatcga cgcgtaccaa cttgccatcc tgaagaatgg 1020

ES 2 661 447 T3

tgcagtgtct cggcacccca tagggaacct ttgccatcaa ctcggaaga tgcagcgtcg	1080
tgttggcatc gtgtcccacg ccgaggagaa gtacctgcc atcgagttca tggacacggg	1140
cgaccgggct tgcaggcgag tgaggtggca ggggcaatgg atcagagatg atctgctctg	1200
cctgtggccc cgctgccgca aaggcaaatg gatgggcgct gcgctttaca tttggcaggc	1260
gccagaatgt gtcagagaca actccaaggt ccggtgtaac gggcgacgtg gcaggatcga	1320
acggctcgtc gtccagacct gaccacgagg gcatgacgag cgtccctccc ggaccacgcg	1380
cagcacgcag ggcctcgatc agtccaagtg gcccatcttc gaggggcccg acgctacgga	1440
aggagctgtg gaccagcagc acaccgccgg gggtaacccc aaggttgaga agctgaccga	1500
tgagctcggc ttttcgcatc tcgtattgca cgacattgca ctccaccgct gatgacatca	1560
gtcgatcata gcacgatcaa cggcactggt gcaaatagtc ggtgggtgata aacttatcat	1620
ccccttttgc ttatggagct gcacatgaac ccattcaaag gccggcattt tcagcgtgac	1680
atcattctgt gggccgtacg ctgggtactgc aaatacggca tcagttaccg tgagctgcat	1740
tttccgctgc ataaccctgc ttcgggggtca ttatagcgat tttttcggtatccatcct	1800
ttttcgacg atatacagga ttttgccaaa gggttcgtgt agactttcct tgggtgatcc	1860
aacggcgtca gccgggcagg ataggtgaag tagggccacc cgcgagcggg tgttccttct	1920
tcactgtccc ttattcgcac ctggcgggtgc tcaacgggaa tcctgctctg cgaggctggc	1980
gggaacttcg aagttcctat actttctaga gaataggaac ttcgaactgc aggtcgacgg	2040
atccccggaa tgtcgacgat gtaggtcacg gtctcgaagc cgcgggtgcgg gtgccagggc	2100
gtgcccttgg gctccccggg cgcgtactcc acctcaccga tctgggtccat catgatgaac	2160
gggtcgaggt ggcggtagtt gatccccggc aacgcgcggc gcaccgggaa gccctcgccc	2220
tcgaaaccgc tgggcgcggg ggtcacgggt agcacgggac gtgcgacggc gtcggcgggt	2280
gcggatacgc ggggcagcgt cagcgggttc tcgacggtca cggcgggcat gtcgaccata	2340
tgaagggcga attcgacca gctttcttgt acaaagttgg cattataaaa aataattgct	2400
catcaatttg ttgcaacgaa caggtcacta tcagtcaaaa taaaatcatt atttgccatc	2460
cagctgatat ccctatagtg gagtcgtatt acatggtcat agctgtttcc tggcagctct	2520
ggcccgtgtc tcaaaatctc tgatgttaca ttgcacaaga taaaaatata tcatcatgcc	2580
tcctctagac cagccaggac agaaatgcct cgacttcgct gctgcccaag gttgccgggt	2640
gacgcacacc gtggaacggt atgaaggcac gaaccagtg gacataagcc tgttcgggtc	2700
gtaagctgta atgcaagtag cgtatgcgct cagcgaactg gtccagaacc ttgaccgaac	2760
gcagcgggtg taacggcgca gtggcgggtt tcatggcttg ttatgactgt ttttttgggg	2820
tacagtctat gcctcgggca tccaagcagc aagcgcgtta cgcggtgggt cgatgtttga	2880
tgttatggag cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca	2940

tcatgagggga agcgggtgatc gccgaagtat cgactcaact atcagaggta gttggcgtca 3000
 tccgagcgcca tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg 3060
 gcggcctgaa gccacacagt gatattgatt tgctggttac ggtgaccgta aggcttgatg 3120
 aaacaacgcg gcgagctttg atcaacgacc ttttggaac ttcggcttcc cctggagaga 3180
 gcgagattct ccgcgctgta gaagtcacca ttgttggtga cgacgacatc attccgtggc 3240
 gttatccagc taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgtag 3300
 gtatcttcga gccagccacg atcgacattg atctggctat cttgctgaca aaagcaagag 3360
 aacatagcgt tgccttggtg ggtccagcgg cggaggaact ctttgatccg gttcctgaac 3420
 aggatctatt tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg 3480
 ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcat ttggtacagc gcagtaaccg 3540
 gcaaaatcgc gccgaaggat gtcgctgcg actgggcaat ggagcgctg ccggcccagt 3600
 atcagcccgat catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg 3660
 cctcgcgcg agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg 3720
 tagtcggcaa ataaccctcg agccacccat gaccaaatac ccttaacgtg agttacgcgt 3780
 cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga gatccttttt 3840
 ttctgcgcgt aatctgctgc ttgcaaacia aaaaaccacc gctaccagcg gtggtttgtt 3900
 tgccggatca agagctacca actctttttc cgaaggtaac tggcttcagc agagcgcaga 3960
 taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag aactctgtag 4020
 caccgcctac atacctcgt ctgctaatac tgttaccagt ggctgctgcc agtggcgata 4080
 agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg 4140
 gctgaacggg gggttcgtgc acacagcca gcttgagcgg aacgacctac accgaactga 4200
 gatacctaca gcgtgagcat tgagaaagcg ccacgcttcc cgaagggaga aagcgggaca 4260
 ggtatccggt aagcggcagg gtcggaacag gagagcgcac gaggagcgtt ccagggggaa 4320
 acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag cgtcgatctt 4380
 tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg gcctttttac 4440
 ggttcctggc cttttgctgg ccttttgcct acatgtt 4477

<210> 10

<211> 58

<212> ADN

5 <213> Artificial

<220>

<223> Cebador 3 directo de PCR

<220>

<221> característica_misc

10 <222> (1)..(39)

<223> Basado en la secuencia de cósmidos

<220>

<221> unión_cebador

<222> (40)..(58)

15 <223> Hibrida con casete aac (3) -attB

<400> 10

tcgacggcct gctgcccgc atcctcaacc agccgacaat gtaggctgga gctgcttc 58

<210> 11

<211> 59

<212> ADN

5 <213> Artificial

<220>

<223> Cebador 3 inverso de PCR

<220>

<221> característica_misc

10 <222> (1)..(39)

<223> Basado en la secuencia de cósmidos

<220>

<221> unión_cebador

<222> (40)..(59)

15 <223> Hibrida con casete aac (3) -attB

<400> 11

aacgcaccgc gcaggcgctg cgcgacggat gggtgaacac atatggtcga catgcccgc 59

<210> 12

<211> 1738

20 <212> ADN

<213> Artificial

<220>

<223> Unión donde el gen de la tiazolinilimida reductasa se eliminó parcialmente dentro del clon del cósmido 4G15

<400> 12

tcgacggcct gctgcccgc atcctcaacc agccgacaat gtaggctgga gctgcttcga 60

agttcctata ttttctagag aataggaact tcggaatagg aacttatgag ctgagccaat 120

cgactggcga gcggcatcgc attcttcgca tcccgccctct ggcggtatgca ggaagatcaa 180

cggatctcgg ccaggttgac ccagggctgt cgccacaatg tcgcggggagc ggatcaaccg 240

agcaaaggca tgaccgactg gaccttcctt ctgaaggctc ttctccttga gccacctgtc 300

cgccaaggca aagcgctcac agcagtggtc attctcgaga taatcgacgc gtaccaactt 360

gccatcctga agaatggtgc agtgtctcgg caccatag ggaacctttg ccatcaactc 420

ggcaagatgc agcgtcgtgt tggcatcgtg tcccacgccg aggagaagta cctgcccac 480

gagttcatgg acacggggcga ccgggcttgc aggcgagtga ggtggcaggg gcaatggatc 540

agagatgatc tgctctgcct gtggccccgc tgccgcaaag gcaaattggat gggcgctgcg 600

25 ctttacattt ggcaggcgcc agaattgtgc agagacaact ccaaggccg gtgtaacggg 660

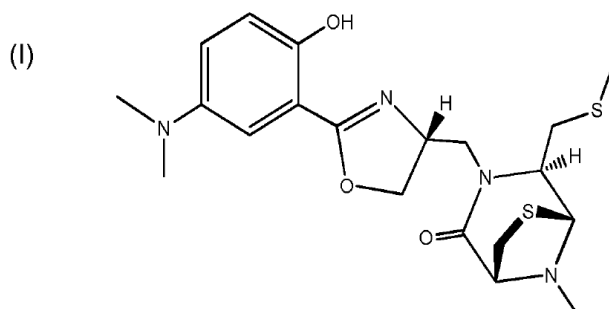
ES 2 661 447 T3

cgacgtggca ggatcgaacg gctcgtcgtc cagacctgac cagaggggca tgacgagcgt	720
ccctcccga cccagcgag cagcgaggc ctcgatcagt ccaagtggcc catcttcgag	780
gggccgggac ctacggaag agctgtggac cagcagcaca ccgccggggg taacccaag	840
ggtgagaagc tgaccgatga gctcggcttt tcgccattcg tattgcacga cattgcactc	900
caccgctgat gacatcagtc gatcatagca cgatcaacgg cactggtgca aatagtcggt	960
ggtgataaac ttatcatccc cttttgctga tggagctgca catgaacca ttcaaaggcc	1020
ggcattttca gcgtgacatc attctgtggg ccgtacgctg gtactgcaa tacggcatca	1080
gttacgtga gctgcatttt ccgctgcata accctgcttc ggggtcatta tagcgatttt	1140
ttcggtatat ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga	1200
ctttccttgg tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gccacccgc	1260
gagcgggtgt tccttcttca ctgtccctta ttcgcacctg gcggtgctca acgggaatcc	1320
tgctctgca ggctggcggg aacttcgaag ttcctatact ttctagagaa taggaacttc	1380
gaactgcagg tcgacggatc cccggaatgt cgacgatgta ggtcacggtc tcgaagccgc	1440
ggtgcgggtg ccagggcgtg cccttgggct ccccgggcgc gtactccacc tcacccatct	1500
ggtccatcat gatgaacggg tcgaggtggc ggtagttgat cccggcgaac gcgcggcgca	1560
ccgggaagcc ctcgccctcg aaaccgctgg gcgcggtggt cacggtgagc acgggacgtg	1620
cgacggcgtc ggcgggtgcg gatacgcggg gcagcgtcag cgggttctcg acggtcacgg	1680
cgggcatgtc gaccatatgt gttcaaccat ccgtcgcgca gcgcctgcgc ggtgcgtt	1738

REIVINDICACIONES

1. Un método para producir una cepa productora de espinosina de *Saccharopolyspora spinosa*; método que comprende:

- 5 modificar una molécula de ácido nucleico que comprende una secuencia de nucleótidos que presenta una identidad de al menos 90% con respecto a SEQ ID NO: 1 mediante mutación de un gen dentro de dicha secuencia de nucleótidos, de tal modo que el gen deja de expresar una proteína funcional; y de tal modo que la cepa mutada produce menos cantidad de un compuesto de Fórmula (I)



que la producida en la cepa antes de la mutación del gen.

- 10 2. El método según la reivindicación 1, en donde la secuencia de nucleótidos es SEQ ID NO: 2.
3. El método según la reivindicación 1, en donde la cepa productora de espinosina produce espinosina J y L.
4. El método según la reivindicación 1, en donde la mutación de la secuencia de nucleótidos comprende la introducción de una eliminación o de un codón de parada.
- 15 5. Una célula de la cepa productora de espinosina de *S. spinosa* obtenida mediante el método de la reivindicación 1 ó 4, en donde la célula comprende una molécula de ácido nucleico que comprende una secuencia de nucleótidos que presenta una identidad de al menos 90% con respecto a SEQ ID NO: 1, y que comprende una mutación en un gen dentro de dicha secuencia de nucleótidos de tal modo que el gen deja de expresar una proteína funcional; y de tal modo que la cepa mutada produce menos cantidad de un compuesto de Fórmula (I) que la producida en una célula antes de la mutación del gen.
- 20 6. El método según la reivindicación 1, que comprende:
- (i) proporcionar una molécula de ácido nucleico que comprende el gen mutado ligado operativamente a al menos una secuencia reguladora de la transcripción; en donde el gen es seleccionado del grupo que consiste en: proteína de unión a sustrato de transportador ABC; permeasa de transportador ABC de aminoácido; transportador de resistencia a fármaco de la familia EmrB QacA; monooxigenasa; NRPS1; NRPS2; un gen localizado entre NRPS1 y NRPS2 en el genoma de *S. spinosa*; tiazolinil imida reductasa; tioesterasa; metiltransferasa; pabAB; un regulador transcripcional; Acil-CoA sintasa; N-metiltransferasa; piridoxamina 5'-fosfato oxidasa; 2,3-dihidrobenczoato-AMP ligasa; y aminotransferasa;
- 25 (ii) introducir la molécula de ácido nucleico en una población de células receptoras de tal modo que la secuencia de nucleótidos se integra de forma estable en el genoma; en donde la célula mantiene la producción de espinosina; y
- 30 (iii) cribar la célula receptora en función de una producción reducida del compuesto de Fórmula (I) en relación con la que se producía en una célula de la misma cepa antes de la mutación del gen.
7. El método según la reivindicación 6, en donde la secuencia de nucleótidos tiene una identidad de al menos 90% con respecto a SEQ ID NO: 2; y en donde la molécula de nucleótido mutada es integrada de forma estable en el ADN genómico del organismo de *S. spinosa* hospedante.
- 35 8. El método según la reivindicación 6, en donde el gen es el gen de *tiazolinil imida reductasa* y la mutación está comprendida en una secuencia que es idéntica en al menos el 80% con respecto a SEQ ID NO: 3.
9. Un organismo según la reivindicación 5, en donde el organismo de *S. spinosa* modificado produce:
- 40 una cantidad de espinosina A que es sustancialmente idéntica a la cantidad de espinosina A producida en el organismo de *S. spinosa* hospedante antes de la introducción del ácido nucleico,

y

una cantidad de espinosina D que es sustancialmente idéntica a la cantidad de espinosina D producida en el organismo de *S. spinosa* hospedante antes de la introducción del ácido nucleico.

10. Un método para producir espinosinas, método que comprende:

5 cultivar el organismo de *Saccharopolyspora spinosa* modificado de la reivindicación 9 en condiciones de fermentación, produciendo con ello al menos una espinosina.

11. Un método según la reivindicación 6, en donde la introducción de la molécula de ácido nucleico en la célula desactiva la producción de un compuesto que tiene la fórmula (I) a la vez que mantiene la producción de espinosina.

FIG. 1.

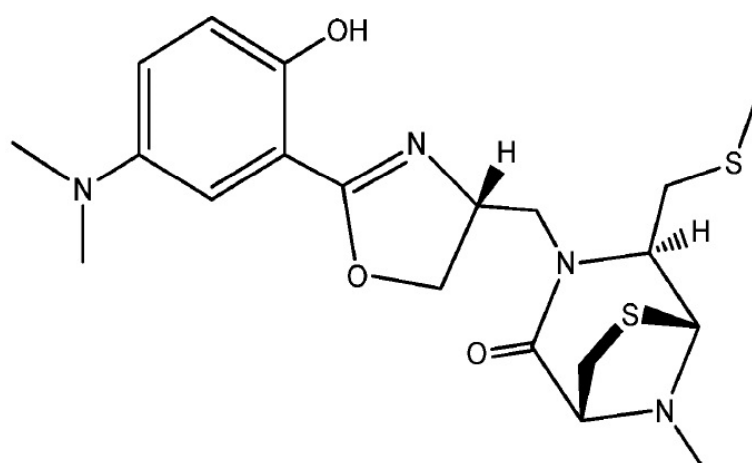


FIG. 2.

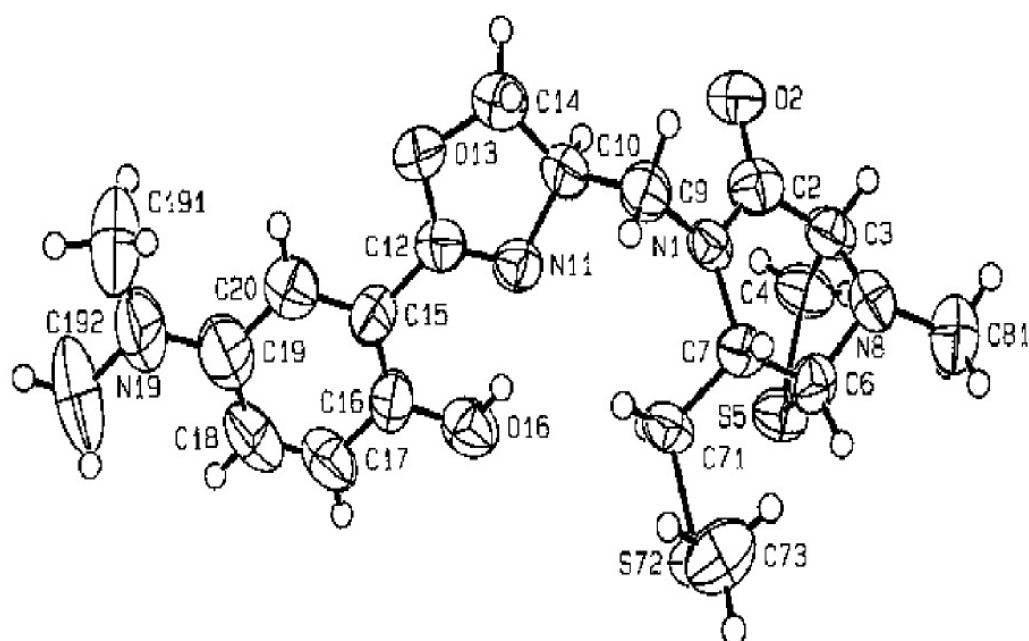


FIG. 3.

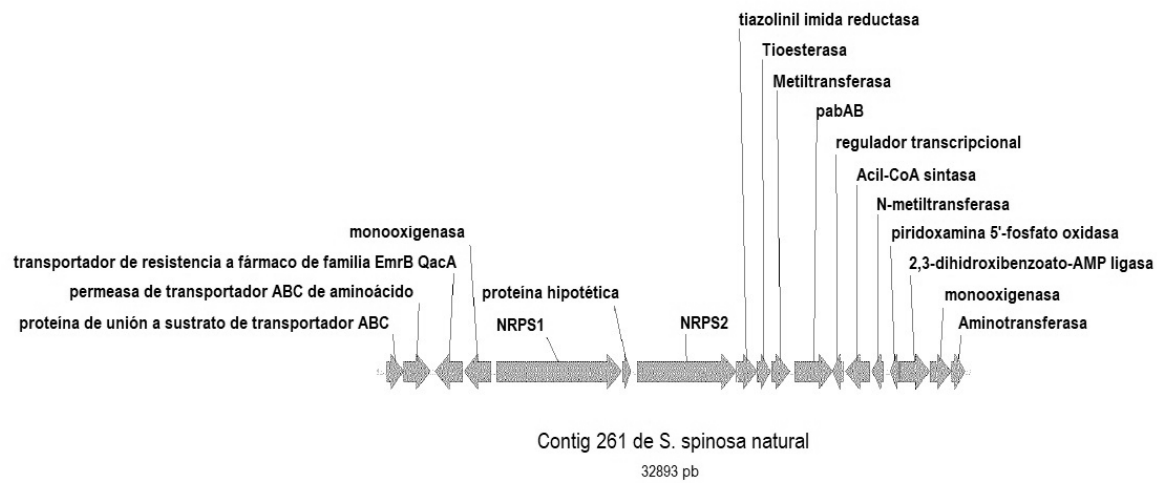


FIG. 4.

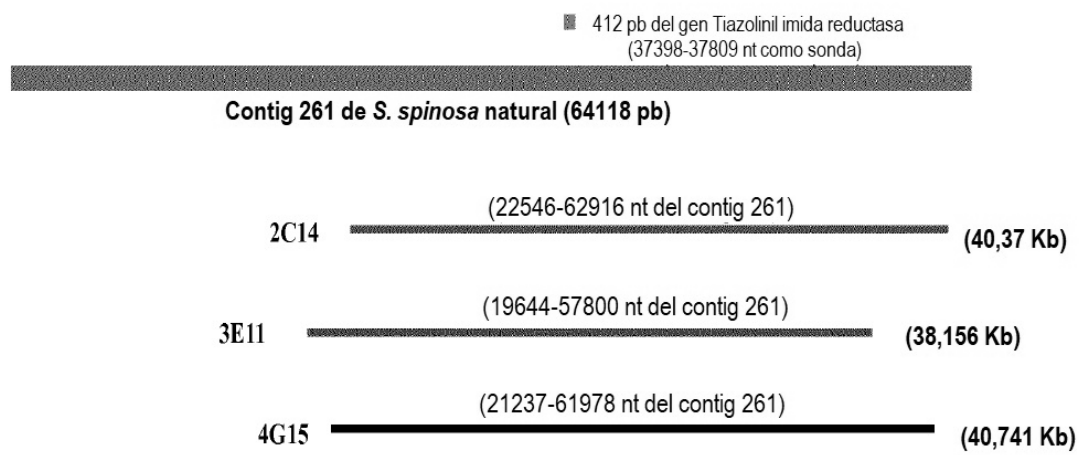


FIG. 5.

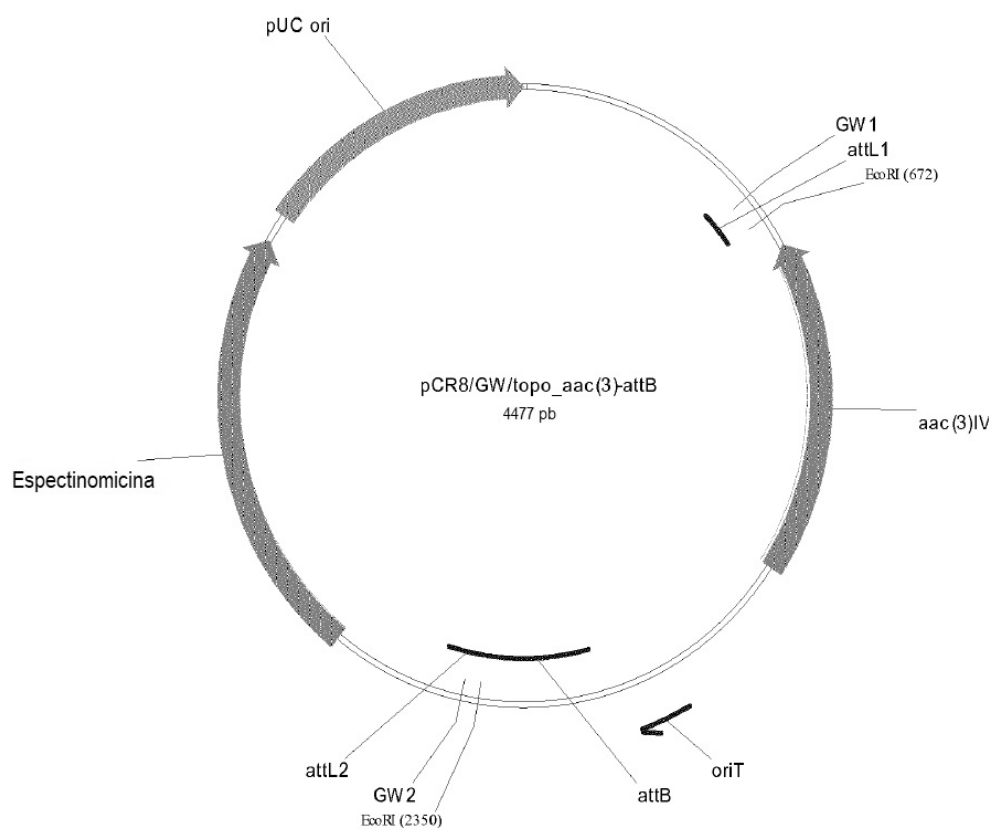


FIG. 6.

