

COMPOSICIONES Y MÉTODOS PARA LA PRODUCCIÓN INCREMENTADA DE ENDURACIDINA EN UNA CEPA GENÉTICAMENTE MODIFICADA DE STREPTOMYCES FUNGICIDICUS

REFERENCIA RECÍPROCA A SOLICITUDES RELACIONADAS

Esta solicitud reivindica el beneficio de prioridad de la fecha de presentación anticipada de las patentes provisionales de EE. UU. Números 62/430.838, presentada el 6 de diciembre de 2016 y 62/479.087, presentada el 30 de marzo de 2017, cada una de las cuales se incorpora aquí como referencia en su totalidad.

CAMPO

Esta descripción se refiere a la biosíntesis de antibióticos, en particular, a las composiciones y métodos para mejorar la producción de enduracidina.

FUNDAMENTO

La aparición mundial de infecciones bacterianas resistentes a múltiples fármacos ha dado lugar a enormes costos sanitarios y se ha convertido en una amenaza importante para la salud pública. Para mantenerse a la vanguardia del desarrollo de resistencias a los medicamentos antibacterianos, existe la necesidad de identificar nuevos antibióticos, así como métodos para producir dichos antibióticos de una manera más rentable.

RESUMEN

La presente divulgación supera los problemas asociados con la producción limitada de enduracidina (enramicina) en cepas de tipo silvestre de *Streptomyces fungicidicus*, así como los límites de producción en cepas industriales desarrolladas a través de la radiación convencional y mutagénesis mediada por químicos del cromosoma y múltiples rondas sucesivas de selección de mutantes para la producción de niveles aumentados de los péptidos deseados del antibiótico enduracidina. En el presente documento se describe la manipulación genética de genes reguladores *orf24* y *orf18* asociados con el grupo de genes de biosíntesis de enduracidina (enramicina) de *Streptomyces fungicidicus* para generar vectores y cepas recombinantes que producen mayores rendimientos de este antibiótico péptidico. Las cepas recombinantes fueron construidas tanto en un productor de tipo silvestre, *Streptomyces fungicidicus* B-5477 (ATCC 21013), y *Streptomyces fungicidicus* BM38-2 (ATCC PTA-122342), que se deriva de la cepa de tipo silvestre y se usa actualmente para la producción

industrial de enduracidina. En el organismo de tipo silvestre, la integración específica a sitio del plásmido pXY152-endorf24, que impulsa la sobreexpresión de una segunda copia de *orf24*, generó la cepa SfpXY152endorf24. La integración del fosmid pXYF24D3 mutagenizado en el cromosoma de tipo Silvestre reemplazó el *orf18* nativo con una copia interrumpida del gen y creó el mutante SfpXYF24D3. Trabajando con el productor commercial de *Streptomyces fungicidicus* BM38-2 (ATCC PTA-122342), la integración del plásmido pXY152-endorf24 generó la cepa recombinante *Streptomyces fungicidicus* BM38-2.24/16. Para crear una cepa derivada de BM38-2 (ATCC PTA-122342) que carece de *orf18* funcional, se construyó el plásmido pKS-T-*orf18pfrd*-AmR para eliminar *orf18* y sus regiones flanqueantes, reemplazando esta región con un marcador de resistencia a apramicina y generando la cepa recombinante de *Streptomyces fungicidicus* BM38-2.18pfrd-AmR. Se demostró que las cepas manipuladas genéticamente producen rendimientos de enduracidina que varían de 1.2 a 4.6 veces más que las cepas precursoras respectivas. Los rendimientos de enduracidina elevados de las cepas recombinantes proporcionan una producción más rentable de enduracidina.

Las anteriores y otras características y ventajas de la divulgación se harán más evidentes a partir de la siguiente descripción detallada, que se desarrolla con referencia a las figuras adjuntas.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La FIG. 1 proporciona la estructura química de las enduracidinas A y B.

La FIG. 2 es un mapa del plásmido integrativo de expresión pXY152-endorf24.

La FIG. 3 es un mapa de la eliminación de genes del plásmido pXY300-*orf18ifd*.

La FIG. 4 es un mapa de la delección de genes del plásmido pKS-T-*orf18ifd*.

La FIG. 5 es un mapa de la delección de genes del plásmido pKS-T-*orf18pfrd*-AmR.

La FIG. 6 es un mapa de la delección de genes del plásmido pKS-*orf18ifd*-T-AmR(NS).

La FIG. 7 es un mapa de la expresión integrativa del plásmido pXY152-endorf24-camtsr.

La FIG. 8 es un mapa de la expresión integrativa del plásmido pXY152-endorf24-blatsr.

La FIG. 9 proporciona una alineación de la proteína StrR activadora de estreptomicina (SEC ID NO: 25) con Orf24 (SEQ ID NO: 26).

Las FIGs. 10A y 10B son mapas de las inserciones de plásmidos pKS-T-orf18pfrd-AmR (a) y pXY300-orf18ifd (b). En el constructo pXY30-orf18ifd, se eliminó y se reemplazó la secuencia interna de *orf18* desde la posición del nucleótido 25795 hasta 26450 (número de acceso de GenBank DQ403252) con un sitio de restricción PacI (TTAATTAA, FIG. 10B). El *orf18* resultante eliminado en el marco (GTGTTTAATTAATGA (SEC ID NO: 27)) podría traducirse en un péptido de tres aminoácidos (VFN). En general, cualquier eliminación interna en el marco a lo largo de *orf18* debería resultar en una función anulada de Orf18 debido a su estado incompleto.

La FIG. 11 es el alineamiento de Orf24 con seis proteínas ortólogas del activador específico de la vía de tipo a StrR funcionalmente caracterizadas de actinomicetos. Orf24 (número de acceso de GenBank DQ403252; SEC ID N°: 26) del grupo de genes biosintéticos de enduracidina de *S. fungicidicus*; StrR (número de acceso de GenBank Y00459; SEC ID N°: 25) del grupo de genes biosintéticos de estreptomicina de *S. griseus*; Tei15* (número de acceso de GenBank AJ632270; SEC ID NO: 32) del grupo de genes de teicoplanina de *Actinoplanes teichomyceticus*; Bbr (número de acceso de GenBank Y16952; SEC ID N°: 28) del grupo de genes biosintéticos *balhimycina* DSM 5908 de la cepa de *Amycolatopsis*; KasT (número de acceso de GenBank BAF79690; SEC ID N°: 29) del grupo de genes *kasugamicina* de *S. kasugaensis*; NovG (número de acceso de GenBank AF170880; SEC ID N°: 30) del grupo de genes biosintéticos de novobiocina NCIMB 9219 de la cepa *S. niveus*; SgcR1 (número de acceso de GenBank AY048670; SEC ID NO: 31) del grupo de genes biosintéticos C-1027 de *S. globisporus*. Aminoácidos idénticos (*), aminoácidos conservadores (.) y sustituciones de aminoácidos altamente conservativos(:). El motivo conservado hélice-giro-hélice (HTH) característico de las proteínas de unión al ADN como StrR está subrayado.

La FIG. 12 Alineación de Orf18 (SEC ID NO: 36) con otros ortólogos reguladores de la respuesta funcionalmente caracterizados. SCO1745/AbrA2: Regulador de respuesta de dos componentes A3 (2) de *S. coelicolor* (número de acceso de GenBank CAB50960; SEC ID NO: 33). SCO3226/AbsA2: Regulador de respuesta de dos componentes A3 (2) de *S. coelicolor* (número de acceso de GenBank AAB08053; SEC ID NO: 34). SCO3818: Regulador transcripcional de respuesta de sistema de dos componentes A3 (2) de *S. coelicolor* (número de acceso de GenBank CAB46941; SEC ID NO: 35).

LISTADO DE SECUENCIAS

Las secuencias de nucleótidos y aminoácidos enumeradas en este documento y en la lista de secuencias adjunta se muestran usando abreviaturas de letras estándar para bases de nucleótidos y un código de tres letras para aminoácidos, como se define en 37 C.F.R. 1.822. Solo se muestra una hebra de la secuencia de ácido nucleico, pero se entiende que la hebra complementaria está incluida por cualquier referencia a la hebra mostrada. En la siguiente lista de secuencias:

Las SEQ ID NOs: 1 y 2 son cebadores oligonucleotídicos utilizados para generar el inserto del plásmido pXY152-endorf24.

La SEQ ID NO: 3 es la secuencia de ácido nucleico del plásmido pXY152-endorf24.

Las SEQ ID NOs: 4-7 son cebadores oligonucleotídicos utilizados para generar el inserto del plásmido pXY300-orf18ifd.

La SEQ ID NO: 8 es la secuencia de ácido nucleico del plásmido pXY300-orf18ifd.

Las SEQ ID NOs: 9 y 10 son cebadores oligonucleotídicos utilizados para generar el fragmento oriT del plásmido pKS-T-orf18pfrd.

La SEQ ID NO: 11 es la secuencia de ácido nucleico del plásmido pKS-T-orf18pfrd.

Las SEQ ID NOs: 12 y 13 son cebadores oligonucleotídicos utilizados para generar el fragmento amR del plásmido pKS-T-orf18pfrd-AmR.

La SEQ ID NO: 14 es la secuencia de ácido nucleico del plásmido pKS-T-orf18pfrd-AmR.

Las SEQ ID NOs: 15-18 son cebadores oligonucleotídicos utilizados para generar los fragmentos oriT y amR del plásmido pKS-orf18ifd-T-AmR (NS).

La SEQ ID NO: 19 es la secuencia de ácido nucleico del plásmido pKS-orf18ifd-T-AmR (NS).

La SEQ ID NO: 20 es la secuencia de ácido nucleico del plásmido pXY152-endorf24-camtsr.

Las SEQ ID NOs: 21 y 22 son cebadores oligonucleotídicos utilizados para generar el fragmento *bla* del plásmido pXY152-endorf24-blatsr.

La SEQ ID NO: 23 es la secuencia de ácido nucleico del plásmido pXY152-endorf24-blatsr.

La SEQ ID NO: 24 es un cebador oligonucleotídico que corresponde a una región de un gen de resistencia a apramicina.

La SEQ ID NO: 25 es la secuencia de aminoácidos de la proteína StrR activadora de estreptomicina.

La SEQ ID NO: 26 es la secuencia de aminoácidos codificada por ORF24.

La SEQ ID NO: 27 es la secuencia de ácido nucleico que ilustra una eliminación en el marco en *orf18*.

La SEQ ID NO: 28 es la secuencia de aminoácidos del inserto Bbr.

La SEQ ID NO: 29 es la secuencia de aminoácidos del inserto KasT.

La SEQ ID NO: 30 es la secuencia de aminoácidos del inserto NovG.

La SEQ ID NO: 31 es la secuencia de aminoácidos del inserto SgcR1.

La SEQ ID NO: 32 es la secuencia de aminoácidos del inserto Teil5*.

La SEQ ID NO: 33 es la secuencia de aminoácidos del ortólogo regulador de la respuesta SCO1745/AbrA2 de *S. coelicolor* A3 (2) (GenBank Accession No. CAB50960).

La SEQ ID NO: 34 es la secuencia de aminoácidos del ortólogo regulador de la respuesta SCO/3226/AbsA2 de *S. coelicolor* A3 (2) (GenBank Accession No. AAB08053).

La SEQ ID NO: 35 es la secuencia de aminoácidos del regulador de respuesta ortólogo SCO3818 de *S. coelicolor* A3 (2) (GenBank Accession No. CAB46941).

La SEQ ID NO: 36 es la secuencia de aminoácidos codificada por ORF18.

La SEQ ID NO: 37 es la secuencia de ácido nucleico de *orf18*.

La SEQ ID NO: 38 es la secuencia de ácido nucleico de *orf24*.

La SEQ ID NO: 39 es la secuencia de ácido nucleico del fosmid pXYF148 con el *orf24* ubicado en la posición de nucleótido 23109 a 24044).

La SEQ ID NO: 40 es la secuencia de ácido nucleico del fosmid pXYF24 con el *orf18* ubicado en la posición del nucleótido 31091-31753).

DESCRIPCIÓN DETALLADA DE VARIAS MODALIDADES

I. Introducción

La enduracidina (FIG. 1), también llamada enramicina, es un antibiótico lipodepsipeptídico de 17 aminoácidos producido por la bacteria del suelo *S. fungicidicus* B-5477 (ATCC 21013). El péptido se aísla del caldo de fermentación y los micelios principalmente como una mezcla de enduracidinas A y B, que difieren en un carbono en la longitud de la cadena lipídica adherida. Estructuralmente, las enduracidinas se distinguen por una fracción de ácido graso ramificado C₁₂ o C₁₃ 2Z, 4E unido por un enlace amida a un residuo de ácido aspártico, y la presencia de numerosos residuos de aminoácidos no proteicos

tales como enduracidina (End), 4-hidroxifenilglicina (Hpg), 3,5-dicloro-4-hidroxifenilglicina (Dpg), citrulina (Cit) y ornitina (Orn) (cf. FIG. 1). Siete de los 17 aminoácidos tienen la configuración D y seis de los residuos son Hpg o el derivado clorado Dpg.

La enduracidina (por simplicidad, se hará referencia a los péptidos de forma singular) muestra una potente actividad antibacteriana in vitro e in vivo contra un amplio espectro de organismos gram-positivos, incluido *Staphylococcus aureus* resistente a meticilina (MRSA) y *Enterococcus* resistente a vancomicina (VRE). Las concentraciones inhibitorias mínimas (CIM) son tan bajas como 0.05 µg/mL y el efecto es bactericida. Un estudio con 100 cepas de *S. aureus* recolectadas de varios productos patológicos, e incluyendo un 40% de MRSA, estableció MIC que varían de 0.09 a 0.56 µg/mL sin una cepa capaz de sobrevivir a la exposición a 1 µg/mL. Para comparación, las CIM típicas para vancomicina hacia cepas sensibles de *S. aureus* varían de 0,5 a 2 µg/ml. Además, la enduracidina tiene un excelente perfil toxicológico. En un estudio en ratones, conejos, perros y monos, las LD50 agudas fueron: intravenosas, 30-125 mg/kg; intraperitoneal, 750-910 mg/kg; subcutánea, intramuscular (i.m.) u oral, > 5-10 g/kg. En el mismo estudio, los monos que recibieron enduracidina por vía intramuscular durante 6 meses y las ratas a las que se les administró una dosis similar durante 12 meses se encontró que solo tenían una inflamación localizada en el sitio de las inyecciones. En humanos, se administró enduracidina por vía intramuscular (100 mg cada 12 horas) a 20 pacientes adultos hospitalizados infectados con MRSA. Se informó que el péptido estaba libre de efectos secundarios y que también era muy efectivo para tratar infecciones del tracto urinario y de la piel causadas por MRSA, pero no infecciones crónicas del hueso (Peromet, Chemotherapy 19: 53-61, 1973).

La enduracidina inhibe la biosíntesis de la pared celular del peptidoglicano bacteriano mediante la formación de complejos con el lípido II extracelular, un precursor para la estructura de la pared celular bacteriana. El sitio de la formación de complejos del lípido II es distinto del reconocido por vancomicina y explica la acción de la enduracidina contra organismos resistentes a vancomicina. Hasta la fecha, no hay una resistencia cruzada documentada de enduracidina con ningún antibiótico utilizado clínicamente y no hay evidencia de resistencia desarrollada, adquirida o transferible. La ausencia de cualquier forma conocida de mecanismo de resistencia transferible, la falta de biodisponibilidad oral, su baja toxicidad y excelente actividad hacia *Clostridium* spp. han hecho de la enduracidina un

antibiótico peptídico comercial clave utilizado como aditivo alimenticio para aves de corral para controlar la enteritis clostridial.

Para derivar una cepa del organismo productor que podría suministrar las cantidades de péptidos requeridos para usos comerciales, Japón Takeda Animal Health (ahora parte de Intervet/Merck Animal Health) sometieron a *S. fungicidicus* B-5477 a varios métodos tradicionales de optimización de cepas y se seleccionaron para los mutantes que produjeron rendimientos más altos de enduracidina. Un mercado mundial creciente de enduracidina ha impulsado los esfuerzos para mejorar aún más el rendimiento de este antibiótico en BM38-2 (ATCC PTA-122342). Con la secuencia genética del grupo de genes de biosíntesis de enduracidina disponible (número de acceso de GenBank DQ403252, que se incorpora aquí como referencia en la red informática mundial el 3 de octubre de 2006, BM38-2 (ATCC PTA-122342) sirvió como la cepa inicial para la manipulación genética dirigida de genes reguladores asociados con el grupo de genes y constituye la base de esta divulgación. En este documento, se describe que el producto de orf18 tiene un efecto negativo en la producción de enduracidina y el producto del gen orf24 tiene un efecto positivo en la producción de enduracidina y que las cepas recombinantes derivadas tanto de *S. fungicidicus* de tipo silvestre como de organismos BM38-2 (ATCC PTA -122342) que explotan estos efectos reguladores producen rendimientos elevados de enduracidina. Además, en el presente documento se describen nuevos vectores de expresión de sustitución génica y de integración basados en pBluescript II KS y pSET152, respectivamente.

II. Abreviaturas y Términos

a. Abreviaturas

AA:	aminoácido
Am:	apramicina
AmR:	marcador de resistencia a apramicina
<i>amRp</i> :	promotor de resistencia a apramicina nativa
ATCC:	Colección de Cultura Tipo Americana
<i>bla</i> :	gen de resistencia a ampicilina
BLAST:	Herramienta básica de búsqueda de alineación local
<i>cam</i> :	gen de resistencia a cloranfenicol
CFU	unidades formadoras de colonias
CTAB:	Bromuro de cetiltrimetilamonio
Cit:	L-citrulina
Dpg:	3,5-dicloro-L-4-hidroxifenilglicina
EDTA:	EtilenDiaminoTetra-Acetato disódico
Final:	enduracidina
Enradin:	Enduracidina, enramicina

EPM:	Medio de producción de enduracidina
Hpg:	D- y L- 4-hidroxifenilglicina
HPLC:	Cromatografía líquida de alto desempeño
HTH:	Hélice-Giro-Hélice
IM:	Intramuscular
ISP2:	Medio 2 del Proyecto Internacional Streptomyces
ISP4:	Medio 4 del Proyecto Internacional Streptomyces
LB:	Caldo Luria-Bertani
LD50:	Dosis letal, un LD50 representa la dosis individual requerida para matar al 50 por ciento de una población de animales de prueba
MAH:	Intervet/Merck Animal Health
MeOH:	Metanol
MICs:	Concentraciones inhibitorias mínimas,
MRSA:	Staphylococcus aureus resistente a meticilina
nm:	Nanómetro
NRPS:	sinetasa de péptidos no ribosomal
ORF:	marco de lectura abierta
Orn:	D-ornitina
PCP:	proteína portadora de peptidilo
PCR:	Reacción en cadena de la polimerasa
Pfrd:	Eliminación adicional de la región flanqueante
SDS:	Dodecilsulfato sódico
SNP:	Polimorfismo de nucleótido único
SPD:	Fotodiodo espectral
TFA:	Ácido trifluoroacético
TSB:	Caldo de soya tríptico
tsr:	Gen de resistencia a tiosstreptona
UV:	ultravioleta
VRE:	enterococos resistentes a vancomicina

b. Condiciones

A menos que se indique lo contrario, los términos técnicos se utilizan de acuerdo con el uso convencional. Las definiciones de términos comunes en biología molecular se pueden encontrar en Benjamin Lewin *Genes V* publicado por Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew, (eds.) *The Encyclopedia of Molecular Biology*, publicado por Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); y Robert A. Meyers (ed.) *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, publicado por VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

Para facilitar la revisión de las diversas modalidades de esta divulgación, se proporcionan las siguientes explicaciones de términos específicos:

Administración: Administración por cualquier vía al animal. Como se usa en este documento, la administración se refiere típicamente a la administración oral.

Variante alélica: Una forma alternativa de un polipéptido que se caracteriza por tener una sustitución, eliminación o adición de uno o más aminoácidos. En un ejemplo, la variante no altera la función biológica del polipéptido.

Amplificación: Cuando se usa en referencia a ácidos nucleicos, técnicas que aumentan el número de copias de una molécula de ácido nucleico en una muestra o espécimen. Un ejemplo de amplificación es la reacción en cadena de la polimerasa, en la cual una muestra biológica recolectada de un sujeto se pone en contacto con un par de cebadores oligonucleotídicos, bajo condiciones que permiten la hibridación de los cebadores con la plantilla de ácido nucleico en la muestra. Los cebadores se extienden bajo condiciones adecuadas, se disocian de la plantilla y luego se vuelven a recocer, se extienden y se disocian para amplificar el número de copias del ácido nucleico. El producto de la amplificación *in vitro* se puede caracterizar por electroforesis, patrones de escisión por endonucleasas de restricción, hibridación o ligación de oligonucleótidos y/o secuenciación de ácidos nucleicos, utilizando técnicas estándar. Otros ejemplos de técnicas de amplificación *in vitro* incluyen la amplificación por desplazamiento de hebras (véase la patente de EE.UU. n° 5.744.311); amplificación isotérmica libre de transcripción (véase la patente de EE.UU. n° 6.033.881); Amplificación de la reacción de reparación de la cadena (véase WO 90/01069); amplificación de la reacción de la cadena ligasa (véase EP-A-320 308); amplificación de la reacción de la cadena de la ligasa para el llenado de faltantes (ver la patente de los Estados Unidos No. 5.427.930); Detección de ligasa acoplada y PCR (véase la patente de EE.UU. No. 6.027.889); y la amplificación libre de transcripción de ARN NASBA™ (véase la patente de EE.UU. n° 6.025.134).

Análogo, derivado o mimético: Un análogo es una molécula que se diferencia en la estructura química de un compuesto precursor, por ejemplo un homólogo (que se diferencia en un incremento en la estructura química, tal como una diferencia en la longitud de una cadena de alquilo), un fragmento molecular, una estructura que difiere por uno o más grupos funcionales, y/o un cambio en la ionización. Los análogos estructurales a menudo se encuentran utilizando relaciones cuantitativas estructura-actividad (QSAR), con técnicas tales como las que se describen en Remington (The Science and Practice of Pharmacology, 19ª edición (1995), capítulo 28). Cuando los cambios al compuesto original son sustanciales, o se combinan muchos cambios incrementales, el compuesto no es más un análogo. Por ejemplo, la ramoplanina no se considera en la presente como un análogo de la enduracidina: la ramoplanina no tiene el aminoácido enduracididina, incluye diferentes aminoácidos, y aunque

tiene una cadena lateral de lípidos, la longitud de cadena es sustancialmente más corta. Los análogos de enduracidina pueden prepararse mediante la adición o eliminación de grupos funcionales en los aminoácidos que constituyen los lipodepsipéptidos, mediante la sustitución de un aminoácido por otro (excepto los aminoácidos enduracididina) o una combinación de modificación del grupo funcional y sustitución de aminoácidos. Ejemplos de análogos de enduracidina incluyen tetrahdoenduracidina A, tetrahydroenduracidina B, descloroenduracidina A y descloroenduracidina B.

Un derivado es una molécula biológicamente activa derivada de la estructura base. Un mimético es una molécula que imita la actividad de otra molécula al imitar la estructura de dicha molécula, tal como una molécula biológicamente activa. Por lo tanto, el término "mimético" indica una estructura definida relacionada con la actividad.

Antibiótico: Una sustancia, por ejemplo, enduracidina, penicilina o estreptomicina, a menudo producida por o derivada de ciertos hongos, bacterias y otros organismos, que puede destruir o inhibir el crecimiento de otros microorganismos.

Antisentido, sentido, y antígenos: El ADN de doble hebra (ADNds) tiene dos hebras, una hebra de 5' → 3', denominada la hebra positiva, y una hebra de 3' → 5' (el complemento inverso), denominada la hebra negativa. Debido a que la ARN polimerasa agrega ácidos nucleicos en una dirección 5' → 3', la hebra negativa del ADN sirve como plantilla para el ARN durante la transcripción. Por lo tanto, el ARN formado tendrá una secuencia complementaria a la hebra negativa e idéntica a la hebra positiva (excepto que U es sustituida por T). Las moléculas antisentido son moléculas que son específicamente hibridables o específicamente complementarias a cualquiera ARN o ADN de hebra positiva. Las moléculas sensoriales son moléculas que son específicamente hibridables o específicamente complementarias a la hebra negativa de ADN. Las moléculas de antígeno son bien sea moléculas complementarias antisentido o sentido a una diana de ADN de doble hebra. En una modalidad, una molécula antisentido hibrida específicamente con un ARNm diana e inhibe la transcripción del ARNm diana.

Unión o unión estable: Una molécula, tal como un oligonucleótido o una proteína, se une o se une de manera estable a una molécula diana, tal como un ácido nucleico o una proteína, si la unión es detectable. En un ejemplo, un oligonucleótido se une o se une de manera estable a un ácido nucleico diana si una cantidad suficiente del oligonucleótido forma pares de bases o se hibrida con su ácido nucleico diana, para permitir la detección de esa unión. La unión puede ser detectada por bien sea las propiedades físicas o funcionales de la

diana: complejo de oligonucleótidos. La unión entre una diana y un oligonucleótido puede detectarse mediante cualquier procedimiento conocido por un experto en la técnica, incluidos los ensayos de unión tanto funcionales como físicos. La unión puede detectarse funcionalmente al determinar si la unión tiene un efecto observable sobre un proceso biosintético tal como la expresión de un gen, replicación del ADN, transcripción, traducción y similares.

Los métodos físicos para detectar la unión de hebras complementarias de ADN o ARN son bien conocidos en la técnica e incluyen métodos tales como DNasa I o huella química, análisis de cambio de gel y ensayos de escisión por afinidad, transferencia Northern, transferencia de puntos y procedimientos de detección de absorción de luz. Por ejemplo, un método que se usa ampliamente, debido a que es tan simple y confiable, implica observar un cambio en la absorción de la luz de una solución que contiene un oligonucleótido (o un análogo) y un ácido nucleico diana en 220 a 300 nm, a medida que la temperatura aumenta poco a poco. Si el oligonucleótido o análogo se ha unido a su diana, hay un aumento repentino en la absorción a una temperatura característica a medida que el oligonucleótido (o análogo) y la diana se disocian entre sí, o se funden.

La unión entre un oligómero y su ácido nucleico diana se caracteriza frecuentemente por la temperatura (T_m) a la que se funde el 50% del oligómero de su diana. Una T_m más alta significa un complejo más fuerte o más estable en relación con un complejo con una T_m más baja.

La unión entre una proteína y su proteína diana, como un anticuerpo para un antígeno, se caracteriza frecuentemente por la determinación de la afinidad de unión. En una modalidad, la afinidad se calcula mediante una modificación del método de Scatchard descrito por Frankel, *Mol. Immunol.*, 16:101-106, 1979. En otra modalidad, la afinidad de unión se mide por una tasa de disociación del receptor de agente de unión específico. En otra modalidad más, una alta afinidad de unión se mide mediante un radioinmunoensayo de competición. En varios ejemplos, una alta afinidad de unión es al menos 1×10^{-8} M aproximadamente. En otras modalidades, una alta afinidad de unión es al menos $1,5 \times 10^{-8}$ aproximadamente, al menos $2,0 \times$

10^{-8} aproximadamente, al menos $2,5 \times 10^{-8}$ aproximadamente, al menos $3,0 \times 10^{-8}$ aproximadamente, al menos $3,5 \times 10^{-8}$ aproximadamente, al menos $4,0 \times 10^{-8}$ aproximadamente, al menos $4,5 \times 10^{-8}$ aproximadamente, o al menos $5,0 \times 10^{-8}$ M aproximadamente.

Función biológica: La función de un polipéptido en las células en las que ocurre naturalmente. Un polipéptido puede tener más de una función biológica.

ADNc (ADN complementario): Una pieza de ADN que carece de segmentos internos no codificantes (intrones) y secuencias reguladoras de la transcripción. El ADNc también puede contener regiones no traducidas (UTR) que son responsables del control de traducción en la molécula de ARN correspondiente. El ADNc se sintetiza en el laboratorio por transcripción inversa del ARN mensajero extraído de las células.

Sustitución conservadora: Las sustituciones de aminoácidos que no alteran sustancialmente la actividad (especificidad o afinidad de unión) de la molécula. Típicamente, las sustituciones de aminoácidos conservadoras implican sustituciones de un aminoácido por otro aminoácido con propiedades químicas similares (ej., carga o hidrofobicidad). La siguiente tabla muestra ejemplos de sustituciones de aminoácidos conservativas:

Residuo original	Sustituciones conservadoras
Ala	Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn; Gln
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; Gln; Glu
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Control de la cepa de *Streptomyces fungicidicus*: La cepa natural de tipo silvestre de origen natural, *Streptomyces fungicidicus* ATCC21013.

ADN (ácido desoxirribonucleico): Un polímero de cadena larga que comprende el material genético de la mayoría de los organismos vivos (algunos virus tienen genes que comprenden el ácido ribonucleico (ARN)). Las unidades de repetición en los polímeros de

ADN son cuatro nucleótidos diferentes, cada uno de los cuales comprende una de las cuatro bases, adenina, guanina, citosina y timina unidas a un azúcar de desoxirribosa al que está unido un grupo fosfato. Las tripletas de nucleótidos (denominadas codones) codifican para cada aminoácido en un polipéptido. El término codón también se usa para las secuencias correspondientes (y complementarias) de tres nucleótidos en el ARNm en el que se transcribe la secuencia de ADN.

A menos que se especifique lo contrario, cualquier referencia a una molécula de ADN pretende incluir el complemento inverso de esa molécula de ADN. Excepto donde se requiere una sola hebra, por el texto en la presente, las moléculas de ADN, aunque están escritas para representar solo una hebra única, abarcan ambas hebras de una molécula de ADN de doble hebra. Por lo tanto, una referencia a la molécula de ácido nucleico que codifica una proteína específica, o un fragmento de la misma, abarca tanto el hilo de los sentidos como su complemento inverso. Así, por ejemplo, es apropiado generar sondas o cebadores a partir de la secuencia de complemento inversa de las moléculas de ácido nucleico descritas.

Dominio: Una porción de una molécula tal como proteínas o ácidos nucleicos que es estructural y/o funcionalmente distinta de otra porción de la molécula.

Codifica: Se dice que un polinucleótido "codifica" un polipéptido si, en su estado nativo o cuando se manipula mediante métodos bien conocidos por los expertos en la técnica, puede transcribirse y/o traducirse para producir el ARNm para y/o el polipéptido o un fragmento del mismo. La hebra antisentido es el complemento de dicho ácido nucleico, y la secuencia de codificación se puede deducir de él.

Enduracidina: Las enduracidinas A y B son lipodepsipéptidos de 17 aminoácidos descubiertos a fines de la década de 1960 a partir de las fermentaciones de la bacteria del suelo *Streptomyces fungicidicus* B-5477 (ATCC 21013). Los péptidos A y B son homólogos que difieren en un carbono en la longitud de la cadena lipídica unida. Estructuralmente, las enduracidinas se distinguen por una fracción de ácido graso ramificado C₁₂ o C₁₃ 2Z, 4E y la presencia de numerosos residuos de aminoácidos no proteicos tales como la enduracididina (End), 4-hidroxifenilglicina (Hpg), 3,5-dicloro-4-hidroxifenilglicina (Dpg), citrulina (Cit) y ornitina (Orn). Siete de los 17 aminoácidos tienen la configuración D y seis de los residuos son Hpg o su análogo clorado Dpg.

Fragmentos funcionales y variantes de un polipéptido: Se incluyen aquellos fragmentos y variantes que mantienen una o más funciones del polipéptido original. Se reconoce que el gen o ADNc que codifica un polipéptido puede mutarse considerablemente

sin alterar materialmente una o más funciones del polipéptido. Primero, el código genético está degenerado y, por lo tanto, diferentes codones codifican los mismos aminoácidos. En segundo lugar, incluso cuando se introduce una sustitución de aminoácidos, la mutación puede ser conservadora y no tener un impacto material en la función o funciones esenciales de una proteína. Véase Stryer Biochemistry, 3ª ed., (C) 1988. Tercero, parte de una cadena polipeptídica puede eliminarse sin afectar o eliminar todas sus funciones. En cuarto lugar, las inserciones o adiciones se pueden hacer en la cadena polipeptídica, por ejemplo, agregando etiquetas de epítipo, sin afectar o eliminar sus funciones (Ausubel, *J. Immunol.* 159(5): 2502-12, 1997). Otras modificaciones que se pueden hacer sin afectar materialmente una o más funciones de un polipéptido incluyen, por ejemplo, modificaciones químicas y bioquímicas in vivo o in vitro o la incorporación de aminoácidos inusuales. Tales modificaciones incluyen, por ejemplo, acetilación, carboxilación, fosforilación, glicosilación, ubiquinación, etiquetado, *ej.*, con radionucleidos, y diversas modificaciones enzimáticas, como apreciarán fácilmente los expertos en la materia. Una variedad de métodos para el etiquetado de polipéptidos, y etiquetas útiles para tales fines, incluyen isótopos radiactivos tales como ³²P, ligandos que se unen o están unidos por parejas de unión específicas etiquetadas (*ej.*, anticuerpos), fluoróforos, agentes quimioluminiscentes, enzimas y antiligandos. Los fragmentos y variantes funcionales pueden ser de longitud variable. Por ejemplo, algunos fragmentos tienen al menos 10, 25, 50, 75, 100, 200 o incluso más residuos de aminoácidos.

Cantidad efectiva: Una cantidad o concentración de un compuesto o composición específica suficiente para lograr un efecto deseado en un sujeto. La cantidad efectiva puede depender, al menos en parte, de la especie de animal que se esté tratando, el tamaño del animal y/o la naturaleza del efecto deseado.

Grupo de genes: Un conjunto de elementos genéticos agrupados en el cromosoma, cuyos productos proteínicos tienen una función relacionada, tal como la formación de una ruta biosintética del producto natural.

Heterólogo: Como se relaciona con secuencias de ácido nucleico tales como secuencias codificantes y secuencias de control, "heterólogo" denota secuencias que normalmente no están asociadas con una región de una construcción recombinante, y/o no están normalmente asociadas con una célula en particular. Por lo tanto, una región "heteróloga" de un constructo de ácido nucleico es un segmento identificable de ácido nucleico dentro de o unido a otra molécula de ácido nucleico que no se encuentra en asociación con la otra molécula en la naturaleza. Por ejemplo, una región heteróloga de un

constructo podría incluir una secuencia de codificación flanqueada por secuencias que no se encuentran en asociación con la secuencia de codificación en la naturaleza. Otro ejemplo de una secuencia de codificación heteróloga es un constructo donde la secuencia de codificación misma no se encuentra en la naturaleza (*ej.*, secuencias sintéticas que tienen codones diferentes al gen nativo). De manera similar, una célula huésped transformada con un constructo que normalmente no está presente en la célula huésped se considera heteróloga para propósitos de esta descripción.

Secuencia homóloga de aminoácidos: Cualquier polipéptido que esté codificado, en su totalidad o en parte, por una secuencia de ácido nucleico que hibrida con cualquier porción de las secuencias de ácido nucleico de la región de codificación. Una secuencia homóloga de aminoácidos es una que difiere de una secuencia de aminoácidos mostrada en la lista de secuencias por una o más sustituciones de aminoácidos conservadoras. Dicha secuencia también abarca variantes alélicas (definidas anteriormente) así como secuencias que contienen deleciones o inserciones que conservan las características funcionales del polipéptido. Preferiblemente, tal secuencia es al menos 75%, más preferiblemente 80%, más preferiblemente 85%, más preferiblemente 90%, más preferiblemente 95%, y más preferiblemente 98% idéntica a cualquiera de las secuencias de aminoácidos.

Las secuencias de aminoácidos homólogas incluyen secuencias que son idénticas o sustancialmente idénticas a las secuencias de aminoácidos de la lista de secuencias. Por "secuencia de aminoácidos sustancialmente idéntica" se entiende una secuencia que es al menos el 90%, preferiblemente 95%, más preferiblemente 97%, y más preferiblemente 99% idéntica a una secuencia de aminoácidos de referencia y que preferiblemente difiere de la secuencia de referencia por una mayoría de sustituciones conservadoras de aminoácidos. De acuerdo con este aspecto de la invención, los polipéptidos que tienen una secuencia homóloga a cualquiera de las secuencias de aminoácidos de la lista de secuencias incluyen variantes alélicas de origen natural, así como mutantes o cualquier otra variante de origen no natural que conserva las características inherentes de cualquier polipéptido de las secuencias descritas en el presente documento. La homología se puede medir utilizando un software de análisis de secuencias como el paquete de software de análisis de secuencia del Genetics Computer Group, Centro de Biotecnología de la Universidad de Wisconsin, 1710 University Avenue, Madison, WI 53705. Las secuencias de aminoácidos pueden alinearse para maximizar la identidad. Los espacios también pueden introducirse artificialmente en la secuencia para lograr una alineación óptima. Una vez que se ha establecido la alineación

óptima, el grado de homología se establece registrando todas las posiciones en las que los aminoácidos de ambas secuencias son idénticos, en relación con el número total de posiciones. Las secuencias de polinucleótidos homólogas se definen de una manera similar. Preferiblemente, una secuencia homóloga es una que es al menos 45%, 50%, 60%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% idéntica a cualquiera de las secuencias de codificación.

Hibridación: Los oligonucleótidos y otros ácidos nucleicos se hibridan mediante enlaces de hidrógeno, que incluyen enlaces de hidrógeno de Watson-Crick, Hoogsteen o Hoogsteen invertidos, entre bases complementarias. En general, el ácido nucleico consiste en bases nitrogenadas que son pirimidinas (citosina (C), uracilo (U) y timina (T)) o purinas (adenina (A) y guanina (G)). Estas bases nitrogenadas forman enlaces de hidrógeno entre una pirimidina y una purina, y la unión de la pirimidina a la purina se conoce como apareamiento de bases. Más específicamente, A se unirá con hidrógeno a T o U, y G se unirá a C. **Complementario** se refiere al apareamiento de bases que se produce entre dos secuencias de ácido nucleico distintas o dos regiones distintas de la misma secuencia de ácido nucleico.

Específicamente hibridables y específicamente complementarios son términos que indican un grado suficiente de complementariedad de tal manera que se produce una unión estable y específica entre un primer ácido nucleico (tal como, un oligonucleótido) y una diana de ADN o ARN. El primer ácido nucleico (tal como, un oligonucleótido) no necesita ser 100% complementario a su secuencia diana para ser específicamente hibridable. Un primer ácido nucleico (tal como, un oligonucleótido) se puede hibridar específicamente cuando existe un grado suficiente de complementariedad para evitar la unión no específica del primer ácido nucleico (tal como, un oligonucleótido) a secuencias no diana bajo condiciones donde la unión específica es deseada. Dicha unión se conoce como hibridación específica.

Las condiciones de hibridación que dan como resultado grados particulares de rigurosidad variarán dependiendo de la naturaleza del método de hibridación elegido y la composición y longitud de las secuencias de ácido nucleico de hibridación. En general, la temperatura de hibridación y la fuerza iónica (especialmente la concentración de Na^+) del tampón de hibridación determinarán la rigurosidad de la hibridación, aunque los tiempos de lavado también influyen en la rigurosidad. Los cálculos con respecto a las condiciones de hibridación requeridas para alcanzar grados particulares de rigurosidad son discutidos por Sambrook (ed.) *Molecular Cloning: A Laboratory Manual*, 2ª ed., Vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, capítulos 9 y 11.

Los siguientes son conjuntos de ejemplo de condiciones de hibridación y no pretenden ser limitantes.

Rigurosidad Muy Alta (detecta secuencias que comparten el 90% de identidad de secuencia)

Hibridización: 5x SSC a 65EC durante 16 horas
Lavar dos veces: 2x SSC a temperatura ambiente (RT) durante 15 minutos cada vez
Lavar dos veces: 0.5x SSC a 65EC durante 20 minutos cada vez

Rigurosidad Alta (detecta secuencias que comparten 80% de identidad de secuencia o más)

Hibridización: 5x-6x SSC a 65EC-70EC durante 16-20 horas
Lavar dos veces: 2x SSC a temperatura ambiente durante 5-20 minutos cada vez
Lavar dos veces: 1x SSC a 55EC-70EC durante 30 minutos cada vez

Rigurosidad Baja (detecta secuencias que comparten más del 50% de identidad de secuencia)

Hibridización: 6x SSC desde temperatura ambiente a 55EC durante 16-20 horas
Lavar al menos dos veces: 2x-3x SSC desde temperatura ambiente a 55EC durante 20-30 minutos cada vez.

Aislado: Un componente biológico aislado (tal como una molécula de ácido nucleico o una proteína) es uno que se ha separado o purificado sustancialmente de otros componentes biológicos en la célula del organismo en el que se encuentra el componente de forma natural, tal como otro ADN cromosómico y extracromosómico y ARN, proteínas y orgánulos. Con respecto a los ácidos nucleicos y/o polipéptidos, el término puede referirse a los ácidos nucleicos o polipéptidos que no se flanquean más por las secuencias que típicamente los flanquean en la naturaleza. Los ácidos nucleicos y proteínas que se han aislado incluyen ácidos nucleicos y proteínas purificadas por métodos de purificación estándar. El término también abarca ácidos nucleicos y proteínas preparados mediante expresión recombinante en una célula huésped, así como ácidos nucleicos sintetizados químicamente.

Mutado: El proceso de realizar un cambio en la secuencia de un material genético (generalmente ADN o ARN) de una célula u organismo. Las mutaciones pueden introducirse intencionalmente en material genético utilizando técnicas moleculares bien conocidas en la técnica (ej., mutagénesis dirigida a sitio, mutagénesis por PCR y otras).

Péptido no ribosomal (PNR): Una clase de metabolitos secundarios, generalmente producidos por microorganismos, tales como bacterias y hongos. A diferencia de los

polipéptidos sintetizados en el ribosoma, estos péptidos se sintetizan mediante sintetasas de péptidos no ribosomales (NRPS) a partir de aminoácidos.

Ensamblaje de la estructura peptídica no ribosomal: El segundo paso en la biosíntesis de péptidos no ribosomales, que incluye la formación de enlaces amida (condensación) de la secuencia peptídica.

Sintetasa de péptidos no ribosomal (NRPS): Una gran proteína multifuncional que sintetiza polipéptidos mediante un mecanismo no ribosomal, a menudo conocido como síntesis de tiotemplato (Kleinkauf y von Doehren *Ann. Rev. Microbiol.* 41: 259-289, 1987). Dichos polipéptidos no ribosomales pueden tener una estructura cíclica lineal, cíclica o ramificada y con frecuencia contienen aminoácidos que no están presentes en proteínas o aminoácidos modificados mediante metilación o epimerización. En ejemplos particulares, la NRPS produce dipéptidos.

Adaptación de péptidos no ribosómicos: El tercer paso en la biosíntesis de péptidos no ribosomales. Hay numerosos aminoácidos precursores nuevos que se encuentran en los péptidos no ribosómicos y muchos de estos precursores se forman o modifican mientras se unen a dominios PCP de proteínas especializadas o la NRPS. Esta modificación post-sintética puede ocurrir después de la formación del enlace amida de la estructura peptídica. Las modificaciones de ejemplo incluyen epimerización de carbono- α , metilación N, heterociclización de residuos de Cys o Ser/Thr a tiazolinas y oxazolinas, y halogenación o hidroxilación de la cadena lateral. Otras modificaciones, como la oxidación, alquilación, acilación y glicosilación, pueden ocurrir después de la liberación del péptido naciente del complejo NRPS y a menudo son necesarias para la actividad biológica completa.

Biosíntesis de aminoácidos precursores no ribosomales: El primer paso en la biosíntesis de péptidos no ribosomales. Los péptidos no ribosomales a menudo poseen aminoácidos que no se encuentran en los péptidos y proteínas que se ensamblan en los ribosomas. Estos aminoácidos no proteicos contribuyen a la diversidad de estos péptidos y, a menudo, desempeñan funciones en su actividad biológica. La biosíntesis de estos aminoácidos puede ocurrir a través de intermedios unidos a proteínas o como especies libres solubles.

Ácido nucleico: Un polímero desoxirribonucleótido o ribonucleótido bien sea en forma de hebra simple o doble, y a menos que se limite de otra manera, abarca análogos conocidos de nucleótidos naturales que se hibridan con ácidos nucleicos de una manera similar a los nucleótidos de origen natural.

Nucleótido: Este término incluye, pero no se limita a, un monómero que incluye una base unida a un azúcar, tal como una pirimidina, purina o análogos sintéticos de los mismos, o una base unida a un aminoácido, como en un ácido nucleico peptídico. Un nucleótido es un monómero en un polinucleótido. Una secuencia de nucleótidos se refiere a la secuencia de bases en un polinucleótido.

Oligonucleótido: Una pluralidad de nucleótidos unidos mediante enlaces fosfodiéster nativos, entre 6 y 300 nucleótidos de longitud aproximadamente. Un análogo de oligonucleótido se refiere a fracciones que funcionan de manera similar a oligonucleótidos pero tienen porciones de origen no natural. Por ejemplo, los análogos de oligonucleótidos pueden contener porciones de origen no natural, tales como fracciones de azúcar alteradas o enlaces entre azúcares, como un oligodesoxinucleótidos fosforotioato. Los análogos funcionales de los polinucleótidos de origen natural pueden unirse al ARN o al ADN, e incluyen moléculas de ácido nucleico peptídico.

Los oligonucleótidos y análogos de oligonucleótidos particulares pueden incluir secuencias lineales de hasta 200 nucleótidos de longitud aproximadamente, por ejemplo, una secuencia (tal como ADN o ARN) que tiene al menos 6 bases, por ejemplo al menos 8, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100 o incluso 200 bases de largo, o de 6 a 50 bases aproximadamente, por ejemplo de 10-25 bases aproximadamente, tales como 12, 15 o 20 bases.

Marco de lectura abierta (ORF): Una serie de tripletes de nucleótidos (codones) que codifican aminoácidos sin ningún codón de terminación interno. Estas secuencias son generalmente traducibles en un péptido. Por ejemplo, ORF, marco de lectura abierta y ORF de enduracidina se refieren a un marco de lectura abierta en el grupo de genes biosintéticos de enduracidina como aislados de *Streptomyces fungicidicus*. El término también abarca los mismos ORF que están presentes en otros organismos que sintetizan la enduracidina. El término abarca variantes alélicas y polimorfismos de un solo nucleótido (SNP). En ciertos casos, el término ORF de enduracidina se usa de forma sinónima con el polipéptido codificado por el ORF de enduracidina y puede incluir sustituciones conservadoras en ese polipéptido. El uso particular será claro desde el contexto.

Un **marco de lectura abierta** que se ha **anulado** es un marco de lectura abierto que se ha vuelto no funcional a través de la eliminación, inserción o mutación de uno o más nucleótidos en la secuencia de codificación.

Un *Streptomyces fungicidicus* que comprende un **marco de lectura abierta disminuido-18 (orf 18)** es un organismo que tiene una disminución en, tal como una disminución de 2 veces, o incluso una pérdida completa de la función biológica del producto génico de *orf18*, en relación a un *Streptomyces fungicidicus* de tipo silvestre, por ejemplo, mediante modificación genética de *orf18*, incluido el *orf18* que se anula como se ejemplifica a continuación, y/o mediante manipulación reguladora, *ej.*, modificación, inserción, eliminación y/o sustitución de regiones no codificantes del gen que codifica ORF18 que resultan en una disminución en la expresión del producto génico *orf18*. Por ejemplo, el promotor de tipo silvestre de *orf18* podría modificarse para disminuir sustancialmente la transcripción de *orf18*.

Un *Streptomyces fungicidicus* que comprende un **marco de lectura abierta aumentada-24 (orf24)** es un organismo que tiene un aumento, por ejemplo un aumento de 2 veces o más, en la función biológica del producto génico de *orf24*, en relación con un *Streptomyces fungicidicus* de tipo silvestre, *ej.*, a través de la modificación genética de *orf24* para mejorar la función biológica del producto génico de *orf24* y/o mediante la manipulación reguladora, *ej.*, modificación, inserción, eliminación y/o reemplazo de regiones no codificantes del gen que codifica ORF24 que resultan en un aumento en la expresión del producto génico *orf24*. Por ejemplo, el promotor de tipo silvestre para *orf24* se reemplazó con un promotor constitutivo fuerte que mejoró la transcripción de *orf24*, como se ilustra a continuación.

Gen modificado: Una secuencia de genes que contiene una modificación en comparación con la que se encuentra en el gen de origen natural (tipo silvestre).

Ligado operativamente: Una primera secuencia de ácido nucleico está unida operativamente con una segunda secuencia de ácido nucleico cuando la primera secuencia de ácido nucleico se coloca en una relación funcional con la segunda secuencia de ácido nucleico. Por ejemplo, un promotor está unido operativamente a una secuencia codificante si el promotor afecta la transcripción o expresión de la secuencia codificante. Generalmente, las secuencias de ADN operativamente unidas pueden ser contiguas y, donde sea necesario unir dos regiones codificantes de proteína, están en el mismo marco de lectura.

Ortólogo: Dos secuencias de ácido nucleico o aminoácido son ortólogas entre sí si comparten una secuencia ancestral común y se separan cuando una especie que lleva esa secuencia ancestral se divide en dos especies. Las secuencias ortólogas también son secuencias homólogas.

Polipéptido: Un polímero en el que los monómeros son residuos de aminoácidos que se unen entre sí mediante enlaces amida. Cuando los aminoácidos son aminoácidos-alfa, se puede usar bien sea el isómero L-óptico o el isómero D-óptico, prefiriéndose los isómeros L en algunos casos. El término polipéptido o proteína como se usa en el presente documento abarca cualquier secuencia de aminoácidos e incluye secuencias modificadas tales como glicoproteínas. El término **polipéptido** se destina específicamente a cubrir las proteínas naturales (ya sean producidas por mecanismos ribosomales o no ribosomales), así como aquellas producidas de forma recombinante o sintética.

El término **fragmento de polipéptido** se refiere a una porción de un polipéptido que exhibe al menos un epítipo útil. La frase **fragmento funcional de un polipéptido** se refiere a todos los fragmentos de un polipéptido que retienen una actividad (tal como una actividad biológica), o una porción medible de una actividad, del polipéptido del cual se deriva el fragmento. Los fragmentos, por ejemplo, pueden variar en tamaño desde un fragmento polipeptídico tan pequeño como un epítipo capaz de unirse a una molécula de anticuerpo a un polipéptido grande capaz de participar en la inducción o programación característica de cambios fenotípicos dentro de una célula.

El término **polipéptido sustancialmente purificado**, como se usa en el presente documento, se refiere a un polipéptido que está sustancialmente libre de otras proteínas, lípidos, carbohidratos u otros materiales con los que está naturalmente asociado. En una modalidad, el polipéptido está al menos en un 50%, por ejemplo en al menos un 80% libre de otras proteínas, lípidos, carbohidratos u otros materiales con los que está asociado de forma natural. En otra modalidad, el polipéptido es al menos en un 90% libre de otras proteínas, lípidos, carbohidratos u otros materiales con los que está naturalmente asociado. En aún otra modalidad, el polipéptido es al menos en un 95% libre de otras proteínas, lípidos, carbohidratos u otros materiales con los que está naturalmente asociado.

Sondas y cebadores: Las sondas de ácido nucleico y cebadores pueden prepararse fácilmente en base a las moléculas de ácido nucleico proporcionadas en esta divulgación. Una sonda comprende un ácido nucleico aislado unido a una etiqueta detectable o molécula reportera. Las etiquetas típicas incluyen isótopos radiactivos, sustratos enzimáticos, co-factores, ligandos, agentes quimioluminiscentes o fluorescentes, haptenos y enzimas. Los métodos para etiquetar y guiar en la elección de etiquetas apropiadas para varios propósitos se discuten, *ej.*, en Sambrook (In *Molecular Cloning: A Laboratory Manual*, CSHL, New

York, 1989) y Ausubel (In *Current Protocols in Molecular Biology*, Greene Publ. Assoc. y Wiley-Intersciences, 1992).

Los cebadores son moléculas de ácido nucleico cortas, preferiblemente oligonucleótidos de ADN, 10 nucleótidos o más de longitud. Más preferiblemente, los oligonucleótidos de ADN más largos pueden tener aproximadamente 15, 17, 20 o 23 nucleótidos o más de longitud. Los cebadores pueden ser recocidos a una hebra de ADN diana complementaria mediante hibridación de ácido nucleico para formar un híbrido entre el cebador y la hebra de ADN diana, y luego el cebador se extendió a lo largo de la cadena de ADN diana por una enzima ADN polimerasa. Los pares de cebadores pueden usarse para amplificar una secuencia de ácido nucleico, *ej.*, mediante la reacción en cadena de la polimerasa (PCR) u otros métodos de amplificación de ácido nucleico conocidos en la técnica.

Los métodos para preparar y usar sondas y cebadores se describen, por ejemplo, en Sambrook (In *Molecular Cloning: A Laboratory Manual*, CSHL, Nueva York, 1989), Ausubel, (In *Current Protocols in Molecular Biology*, Greene Publ. Assoc. y Wiley-Intersciences, 1998), e Innis, (*PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc., San Diego, CA, 1990). Los pares de cebadores de PCR se pueden derivar de una secuencia conocida, por ejemplo, mediante el uso de programas informáticos destinados a tal fin, como Primer (Versión 0.5, © 1991, Instituto Whitehead para la Investigación Biomédica, Cambridge, MA).. La especificidad de una sonda o cebador particular aumenta con su longitud. Por lo tanto, para obtener una mayor especificidad, se pueden seleccionar sondas y cebadores que comprenden al menos 17, 20, 23, 25, 30, 35, 40, 45, 50 o más nucleótidos consecutivos de la secuencia de nucleótidos deseada.

Proteína: Una molécula biológica expresada mediante un gen y compuesta de aminoácidos.

Purificado: El término **purificado** no requiere pureza absoluta; más bien, se pretende como un término relativo. Así, por ejemplo, una preparación de proteína purificada es aquella en la que la proteína a la que se hace referencia es más pura que la proteína en su entorno natural dentro de una célula.

Recombinante: Un ácido nucleico que tiene una secuencia que no se produce de forma natural o tiene una secuencia creada por una combinación artificial de dos segmentos de secuencia separados de otra manera. Esta combinación artificial puede lograrse mediante la síntesis química o, más comúnmente, mediante la manipulación artificial de segmentos

aislados de ácidos nucleicos, *ej.*, mediante técnicas de ingeniería genética. "Recombinante" también se usa para describir moléculas de ácido nucleico que han sido manipuladas artificialmente, pero contienen las mismas secuencias de control y regiones codificantes que se encuentran en el organismo del cual se aisló el gen.

Regulación de la producción de antibióticos: Para causar una alteración, tal como un aumento o disminución, en la cantidad, tipo o calidad de la producción de antibióticos. En el presente documento se describen cepas recombinantes de *Streptomyces fungicidicus* con producción de enduracidina mejorada.

Identidad de secuencia: La similitud entre dos secuencias de ácido nucleico o entre dos secuencias de aminoácidos se expresa en términos de nivel de identidad de secuencia compartida entre las secuencias. La identidad de secuencia se expresa típicamente en términos de porcentaje de identidad; cuanto mayor sea el porcentaje, más similares serán las dos secuencias.

Los métodos para alinear secuencias para comparación son bien conocidos en la técnica. Varios programas y algoritmos de alineación se describen en: Smith y Waterman, *Adv. Appl. Math.* 2:482, 1981; Needleman y Wunsch, *J. Mol. Biol.* 48:443, 1970; Pearson y Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444, 1988; Higgins y Sharp, *Gene* 73:237-244, 1988; Higgins y Sharp, *CABIOS* 5:151-153, 1989; Corpet, *Nucleic Acids Research* 16:10881-10890, 1988; Huang, *Computer Applications in the Biosciences* 8:155-165, 1992; Pearson, *Methods in Molecular Biology* 24:307-331, 1994; Tatiana, (1999), *FEMS Microbiol. Lett.*, 174:247-250, 1999. Altschul, presentó una consideración detallada de los métodos de alineación de secuencias y cálculos de homología (*J. Mol. Biol.* 215:403-410, 1990).

El Centro Nacional de Información Biotecnológica (NCBI) Herramienta de Búsqueda Básica de Alineación Local (BLASTTM, Altschul. *J. Mol. Biol.* 215: 403-410, 1990) está disponible en varias fuentes, incluido el Centro Nacional de Información Biotecnológica (NCBI, Bethesda, MD) y en Internet, para su uso en relación con los programas de análisis de secuencias blastp, blastn, blastx, tblastn y tblastx. Una descripción de cómo determinar la identidad de secuencia mediante este programa está disponible en Internet en la sección de ayuda de BLASTTM.

Para las comparaciones de secuencias de aminoácidos de más de 30 aminoácidos aproximadamente, la función "Secuencias Blast 2" del programa BLASTTM (Blastp) se emplea utilizando la matriz BLOSUM62 predeterminada establecida en los parámetros predeterminados (costo de abrir un espacio[predeterminado = 5]; costo de ampliar un espacio

[predeterminado = 2]; penalización por una falta de coincidencia [predeterminado = 3]; recompensa por una coincidencia [predeterminado = 1]; valor de expectativa (E) [predeterminado = 10.0]; tamaño de palabra [predeterminado = 3]; número de descripciones de una línea (V) [predeterminado = 100]; número de alineaciones a mostrar (B) [predeterminado = 100]). Al alinear péptidos cortos (menos de 30 aminoácidos aproximadamente), la alineación debe realizarse utilizando la función de secuencias Blast 2, empleando la matriz PAM30 ajustada a los parámetros predeterminados (brecha abierta 9, penalizaciones de la brecha de extensión 1). Las proteínas (o ácidos nucleicos) con aún mayor similitud con las secuencias de referencia mostrarán identidades porcentuales mayores cuando se evalúen mediante este método, como por ejemplo al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 85 %, al menos 90%, o al menos 95% de identidad de secuencia.

Para comparaciones de secuencias de ácido nucleico, la función de "secuencias Blast 2" del programa BLAST TM (Blastn) se emplea mediante el uso de la matriz BLOSUM62 predeterminada configurada a los parámetros predeterminados (costo para abrir una brecha [por defecto = 11]; costo para ampliar una brecha [por defecto = 1]; valor de expectativa (E) [por defecto = 10.0]; tamaño de palabra [por defecto= 11]; número de descripciones de una línea (V) [por defecto = 100]; número de alineaciones a mostrar (B) [por defecto = 100]). Secuencias de ácidos nucleicos con aún mayor similitud con las secuencias de referencia mostrarán identidades porcentuales mayores cuando se evalúen mediante este método, como por ejemplo al menos 60%, al menos 70%, al menos 75%, al menos 80%, al menos 85 %, al menos 90%, al menos 95%, o al menos 98% de identidad de secuencia.

Una indicación alternativa de que dos moléculas de ácido nucleico están estrechamente relacionadas es que las dos moléculas se hibridan entre sí en condiciones rigurosas (véase "Hibridación" más arriba).

Sin embargo, las secuencias de ácidos nucleicos que no muestran un alto grado de identidad pueden codificar secuencias de aminoácidos similares, debido a la degeneración del código genético. Se entiende que los cambios en la secuencia de ácido nucleico se pueden hacer usando esta degeneración para producir múltiples moléculas de ácido nucleico que todas codifican sustancialmente la misma proteína.

Transfectado: Un proceso por el cual una molécula de ácido nucleico se introduce en la célula, por ejemplo, mediante técnicas de biología molecular, que resultan en **una célula transfectada (o transformada)**. Tal como se usa en el presente documento, el término

transfección abarca todas las técnicas mediante las cuales se podría introducir una molécula de ácido nucleico en dicha célula, incluida la transducción con vectores virales, transfección con vectores plasmídicos y introducción de ADN por electroporación, lipofección y pistola de aceleración de partículas.

Transformado: Una célula transformada es una célula en la que se ha introducido una molécula de ácido nucleico mediante técnicas de biología molecular. El término abarca todas las técnicas mediante las cuales se puede introducir una molécula de ácido nucleico en dicha célula, incluida la transfección con vectores virales, la transformación con vectores plásmidos y la introducción de ADN desnudo mediante electroporación, lipofección y pistola de aceleración de partículas.

Transposón: Un elemento genético móvil que tiene secuencias repetitivas casi idénticas en cada extremo, y que contiene al menos un gen que codifica una transposasa (la enzima necesaria para insertar el transposón en la secuencia de ADN). Los transposones se pueden integrar en diferentes posiciones en el genoma de una célula, o sobre un plásmido aislado, cósmido o plantilla de ADN de fosmid in vitro. Los transposones también pueden contener genes distintos de los necesarios para la inserción.

Vector: Una molécula de ácido nucleico introducida en una célula huésped, produciendo así una célula huésped transfectada. Los vectores de ADN recombinante son vectores que tienen ADN recombinante. Un vector puede incluir secuencias de ácido nucleico que le permiten replicarse en una célula huésped, tal como un origen de replicación. Un vector también puede incluir uno o más genes marcadores seleccionables y otros elementos genéticos conocidos en la técnica. Los vectores virales son vectores de ADN recombinante que tienen al menos algunas secuencias de ácido nucleico derivadas de uno o más virus. Un plásmido es un vector.

A menos que se explique lo contrario, todos los términos técnicos y científicos usados en la presente tienen el mismo significado que entiende comúnmente un experto habitual en la técnica a la que pertenece esta descripción. Los términos singulares "un", "una" y "el" incluyen referentes plurales, a menos que el contexto indique claramente lo contrario. De manera similar, la palabra "o" pretende incluir "y" a menos que el contexto indique claramente lo contrario. Además, debe entenderse que todos los tamaños de bases o tamaños de aminoácidos, y todos los valores de peso molecular o masa molecular, dados para los ácidos nucleicos o polipéptidos son aproximados, y se proporcionan para como descripción.

Aunque los métodos y materiales similares o equivalentes a los descritos en la presente pueden ser utilizados en la práctica o ensayo de esta divulgación, los métodos y materiales adecuados se describen a continuación. El término "comprende" significa "incluye". En caso de conflicto, prevalecerá la presente especificación, incluidas las explicaciones de los términos. Adicionalmente, los materiales, métodos y ejemplos son sólo ilustrativos y no pretenden ser limitantes.

Los métodos y materiales adecuados para la práctica de las modalidades descritas se describen a continuación. Además, cualquier método o técnica apropiada bien conocida por los expertos en la técnica puede usarse en la realización de las modalidades descritas. Algunos métodos y técnicas convencionales aplicables a la presente divulgación se describen, por ejemplo, en Sambrook, *Molecular Cloning: A Laboratory Manual*, 2^{da} ed., Cold Spring Harbor Laboratory Press, 1989; Sambrook, *Molecular Cloning: A Laboratory Manual*, 3ra edición, Cold Spring Harbor Press, 2001; Ausubel, *Current Protocols in Molecular Biology*, Greene Publishing Associates, 1992 (y Supplements to 2000); Ausubel, *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology*, 4ta edición., Wiley & Sons, 1999; Harlow y Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 1990; Harlow y Lane, *Using Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 1999; y Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., y Hopwood, D.A.: *Practical Streptomyces genetics*, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 &UH, England, 2000.

Todas las publicaciones, solicitudes de patentes, patentes, y otras referencias mencionadas en el presente documento se incorporan como referencia en su totalidad. En caso de conflicto, prevalecerá la presente especificación, incluidas las explicaciones de los términos. Adicionalmente, los materiales, métodos y ejemplos son sólo ilustrativos y no pretenden ser limitantes.

III. Vectores de expresión recombinantes modificados de *Streptomyces fungicidicus*

En el presente documento se describen vectores plasmídicos de expresión de *Streptomyces fungicidicus* recombinantes modificados. En algunas modalidades, un vector de *Streptomyces fungicidicus* recombinante modificado comprende al menos un marco de lectura abierta seleccionado de *Streptomyces fungicidicus*. En algunas modalidades, un vector de *Streptomyces fungicidicus* recombinante modificado comprende al menos un marco de lectura abierta seleccionado de *Streptomyces fungicidicus* expresado bajo el control de un

promotor. En algunos ejemplos, el promotor es un promotor de *Streptomyces* constitutivo fuerte que da como resultado la producción mejorada de enduracidina cuando el vector se expresa en una cepa de *Streptomyces fungicidicus*. En algunas modalidades, el marco de lectura abierta está unido operativamente a un promotor heterólogo en lugar de su propio promotor nativo. Por ejemplo, puede estar operativamente unido a un promotor constitutivo, tal como un promotor de expresión constitutivo fuerte o un promotor inducible. En algunos ejemplos, el promotor constitutivo fuerte es ermE*p del productor de eritromicina. En algunos ejemplos, el promotor inducible es tipA. En algunos ejemplos, el sistema P(nitA)-NitR (Herai S, Hashimoto Y, Higashibata H, Maseda H, Ikeda H, Omura S, Kobayashi M, Proc Natl Acad Sci U S A.2004. 101(39): 14031-5) o se emplea el promotor de estreptomicetos SF14. En algunos ejemplos, se emplea un promotor nativo del gen resistente a la apramicina (amRp). En algunos ejemplos, se emplean P_{hrdB}, P_{tcp830}, P_{SF14}, P_{ermE*} y/o P_{neos}.

En algunas modalidades, el vector recombinante modificado comprende un marco de lectura abierta orf 24 (SEC ID NO: 38) y/o marco de lectura abierta orf18 (SEC ID NO: 37) el cual ha sido anulado. En algunos ejemplos, el marco de lectura abierta orf18 (SEQ ID NO: 37) se anula mediante una eliminación en marco, cambio de marco y/o mutación puntual.

En algunas modalidades, el vector recombinante modificado comprende un marco de lectura abierta orf24 del grupo de genes de enduracidina de *Streptomyces fungicidicus*. En algunos ejemplos, el marco de lectura abierta orf24 (SEQ ID NO: 38) está unido operativamente a un promotor heterólogo. Por ejemplo, está vinculado a un promotor constitutivo fuerte como ermE*p. En otros ejemplos, el marco de lectura abierta orf24 está enlazado operativamente al promotor tipA, SF14, amRp, P_{hrdB}, P_{tcp830}, P_{SF14}, P_{ermE*} y/o P_{neos}.

En otra modalidad, un vector recombinante modificado comprende un marco de lectura abierta orf18 que reside en la región corriente arriba del grupo de genes de enduracidina. El marco de lectura abierta orf18 (SEQ ID NO: 37) se anula por interrupción de inserción, eliminación en marco, cambio de marco y/o mutación puntual. En algunos ejemplos, el marco de lectura abierto orf18 se anula mediante una eliminación en marco, tal como una eliminación en marco como se ilustra en la FIG. 9B. En un ejemplo, el marco de lectura abierta orf18 (SEQ ID NO: 37) se anula mediante una eliminación en el marco. Por ejemplo, el marco de lectura abierta orf18 (SEQ ID NO: 37) se anula mediante una eliminación en el marco de los ácidos nucleicos 5 a 660 de orf18 (SEC ID NO: 37). En general, cualquier eliminación interna en marco sobre orf18 da como resultado una función

anulada de Orf18 debido a su estado incompleto. En algunos ejemplos, la eliminación en marco incluye la eliminación de al menos 3 ácidos nucleicos en *orf18* (SEC ID NO: 37), tales como al menos 3 ácidos nucleicos, incluyendo 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87, 90, 93, 96, 99, 102, 105, 108, 111, 114, 117, 120, 123, 126, 129, 132, 135, 138, 141, 144, 147, 150, 153, 156, 159, 162, 165, 168, 171, 174, 177, 180, 183, 186, 189, 192, 195, 198, 201, 204, 207, 210, 213, 216, 219, 221, 224, 227, 230, 233, 236, 239, 245, 248, 251, 254, 257, 260, 263, 266, 269, 272, 275, 278, 281, 284, 287, 290, 293, 296, 299, 302, 305, 308, 311, 314, 317, 320, 323, 326, 329, 332, 335, 341, 344, 347, 350, 353, 356, 359, 362, 365, 368, 371, 374, 377, 380, 383, 386, 389, 392, 398, 401, 404, 407, 410, 413, 416, 419, 421, 424, 427, 430, 433, 436, 439, 442, 445, 448, 451, 454, 457, 460, 463, 466, 469, 472, 475, 478, 484, 487, 490, 493, 496, 499, 502, 505, 508, 511, 514, 517, 520, 523, 526, 529, 532, 535, 538, 544, 547, 550, 553, 556, 552, 562, 565, 568, 571, 574, 577, 580, 583, 586, 589, 592, 595, 598, 601, 604, 607, 610, 613, 616, 619, 621, 627, 630, 633, 636, 639, 642, 645, 648, 651, o 654 ácidos nucleicos entre los ácidos nucleicos 5 a 660 de *orf18* (SEC ID NO: 37).

En modalidades relacionadas, un vector de plásmido recombinante modificado involucra dos o más marcos de lectura abierta del grupo de genes de enduracidina y/o las regiones que flanquean el grupo de genes o de otras cepas de actinomicetos. Los dos o más marcos de lectura abierta pueden estar vinculados a un único promotor. Alternativamente, pueden estar vinculados operativamente a dos promotores diferentes. Los dos promotores pueden ser del mismo tipo de promotor. Alternativamente, pueden ser dos tipos de promotores diferentes.

En modalidades adicionales, se pueden introducir, o inactivar, marcos de lectura abierta adicionales o alternativos que pueden mejorar la producción de enduracidina en la cepa modificada de *Streptomyces fungicidicus*.

En algunos ejemplos, el plásmido recombinante es pXY152-endorf24 (SEC ID NO: 3). En algunos ejemplos, el plásmido recombinante es pXY300-orf18ifd (SEC ID NO: 8). En algunos ejemplos, el plásmido recombinante es pKS-T-orf18ifd (SEC ID NO: 11). En algunos ejemplos, el plásmido recombinante es pKS-T-orf18pfrd-AmR (SEC ID NO: 14). En algunos ejemplos, el plásmido recombinante es pKS-orf18ifd-T-AmR (NS) (SEC ID NO: 19). En algunos ejemplos, el plásmido recombinante es pXY152-endorf24-camtsr (SEC ID NO: 20). En algunos ejemplos, el plásmido recombinante es pXY152-endorf24-blatsr (SEC ID NO: 23).

IV. Cepas recombinantes modificadas de *Streptomyces fungicidicus*

En el presente documento se describen cepas de *Streptomyces fungicidicus* recombinantes modificadas para producir enduracidina mejorada en comparación con una cepa de control (tal como una cepa de *Streptomyces fungicidicus* tipo silvestre o una cepa parental industrial). En algunas modalidades, una cepa de *Streptomyces fungicidicus* recombinante modificada comprende al menos un marco de lectura abierta seleccionado de *Streptomyces fungicidicus* introducido en el cromosoma y expresado bajo el control de un promotor, tal como un promotor de *Streptomyces* constitutivo fuerte, que resulta en la producción mejorada de enduracidina en la cepa modificada. En algunas modalidades, la expresión del marco de lectura abierta introducido en el *Streptomyces fungicidicus* es dirigida por un promotor heterólogo en lugar de su propio promotor nativo. Por ejemplo, puede estar operativamente unido a un promotor constitutivo, tal como un promotor de expresión constitutivo fuerte o un promotor inducible. En algunos ejemplos, el promotor constitutivo fuerte es *ermE**p del productor de eritromicina. En algunos ejemplos, el promotor inducible es *tipA*. En algunos ejemplos, el sistema P(*nitA*)-NitR (véase, Herai S, Hashimoto Y, Higashibata H, Maseda H, Ikeda H, Omura S, Kobayashi M, Proc Natl Acad Sci U S A.2004. 101(39): 14031-5) o se emplea el promotor de estreptomicetos SF14. En algunos ejemplos, el promotor de expresión constitutivo es *amRp*. En algunos ejemplos, se emplean los promotores *P_{hrdB}*, *P_{tcp830}*, *P_{SF14}*, *P_{ermE}** y/o *Pneos*.

En algunas modalidades, la cepa modificada comprende un marco de lectura abierta *orf24* del grupo de genes de enduracidina de *Streptomyces fungicidicus*. En algunos ejemplos, el marco de lectura abierta *orf24* está unido operativamente a un promotor heterólogo. Por ejemplo, está vinculado a un promotor constitutivo fuerte como *ermE**p. En otros ejemplos, el marco de lectura abierta *orf24* está enlazado operativamente al promotor *tipA*, SF14, *amRp*, *P_{hrdB}*, *P_{tcp830}*, *P_{SF14}*, *P_{ermE}** y/o *Pneos*.

En otra modalidad, la cepa modificada se relaciona con un marco de lectura abierta *orf18* que reside en la región corriente arriba del grupo de genes de enduracidina. El marco de lectura abierta *orf18* se anula por interrupción de inserción, eliminación en marco, desplazamiento de marco y/o mutación de punto. En algunos ejemplos, el marco de lectura abierto *orf18* se anula mediante una eliminación en marco, tal como una eliminación en marco como se ilustra en la FIG. 9B. En un ejemplo, el marco de lectura abierta *orf18* (SEQ ID NO: 37) se anula mediante una eliminación en el marco. Por ejemplo, el marco de lectura abierta *orf18* (SEQ ID NO: 37) se anula mediante una eliminación en el marco de los ácidos

nucleicos 5 a 660 de *orf18* (SEC ID NO: 37). En general, cualquier eliminación interna en marco sobre *orf18* debería resultar en una función anulada de Orf18 debido a su estado incompleto.

En modalidades relacionadas, la cepa modificada implica dos o más marcos de lectura abierta del grupo de genes de enduracidina y/o las regiones que flanquean el grupo de genes o de otras cepas de actinomicetos. Los dos o más marcos de lectura abierta pueden estar vinculados a un único promotor. Alternativamente, pueden estar vinculados operativamente a dos promotores diferentes. Los dos promotores pueden ser del mismo tipo de promotor. Alternativamente, pueden ser dos tipos de promotores diferentes.

En modalidades adicionales, se pueden introducir, o inactivar, marcos de lectura abierta adicionales o alternativos que pueden mejorar la producción de enduracidina en la cepa modificada de *Streptomyces fungicidicus*.

En algunas modalidades, la cepa modificada de *Streptomyces fungicidicus* se deriva de una cepa precursora tipo silvestre, tal como, pero sin limitarse a, *Streptomyces fungicidicus*, American Tissue Culture Company (ATCC) 21013. En otras modalidades, la cepa modificada de *Streptomyces fungicidicus* se deriva de una cepa precursora industrial, tal como, pero sin limitarse a, BM38-2 (ATCC PTA-122342). En otras modalidades, la cepa modificada de *Streptomyces fungicidicus* se deriva de las cepas mutantes convencionales, tales como, pero sin limitarse a *Streptomyces fungicidicus* ATCC 31729, *Streptomyces fungicidicus* ATCC 31730 y *Streptomyces fungicidicus* ATCC 31731.

En algunas modalidades, la producción mejorada de enduracidina es un aumento de al menos 1,2 veces, tal como al menos 1,5 veces, al menos 2 veces, al menos 2,5 veces, al menos 3 veces, al menos 3,5 veces, al menos 4 veces, al menos un aumento de 4.5 veces, que incluye, pero no se limita a, un aumento de 1.2 a 10 veces, un aumento de 1.2 a 4.6 veces, un aumento de 2 a 5 veces, tal como 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, y 10 veces más en la producción de enduracidina en comparación con la cepa de control de *Streptomyces fungicidicus*. En algunas modalidades, la cepa de *Streptomyces fungicidicus* de control es una cepa de *Streptomyces fungicidicus* de tipo silvestre, que incluye, entre otros, la Compañía Americana de Cultivos de Tejidos de *Streptomyces fungicidicus* (ATCC) 21013 o una cepa industrial precursora, como por ejemplo, pero no limitada a, BM38 -2 (ATCC PTA-122342), o la cepa mutante convencional, tal como, pero no limitada a *Streptomyces fungicidicus* ATCC 31729,

Streptomyces fungicidicus ATCC 31730 y *Streptomyces fungicicicus* ATCC 31731. En un ejemplo, el control es *Streptomyces fungicidicus* ATCC 21013 y el aumento en la producción mejorada de enduracidina es al menos un aumento de 1.2 veces, tal como un aumento de 1.2 a 4.6 veces. En un ejemplo, el control es *Streptomyces fungicidicus* ATCC 21013 y el aumento en la producción mejorada de enduracidina es al menos un aumento de 1.2 veces, como un aumento de 1.2 a 4.6 veces.

V. Construcción de cepas recombinantes modificadas de Streptomyces fungicidicus

En modalidades, las cepas recombinantes de *Streptomyces fungicidicus* pueden construirse mediante la integración de un plásmido recombinante que comprende al menos un marco de lectura abierto que mejora la producción de enduracidina en el cromosoma de una cepa precursora de *Streptomyces fungicidicus*. El vector conyugal integrativo puede tener, o puede modificarse para tener, un promotor de constitutivo fuerte de *Streptomyces*. En algunas modalidades, el plásmido puede carecer de un replicón de estreptomicetos y puede integrarse en el cromosoma por recombinación homóloga de cruce único específico del sitio. En otras modalidades, el plásmido puede estar presente como un plásmido libre. En algunas modalidades, un vector conyugal puede modificarse en el que el inserto de plásmido porta un gen de interés parcial o completamente eliminado, y sus regiones flanqueantes, que pueden integrarse en el cromosoma después de una doble recombinación homóloga cruzada para generar un mutante con eliminación en el marco.

VI Producción de enduracidina a partir de cepas recombinantes modificadas de Streptomyces fungicidicus

Las cepas recombinantes modificadas de *Streptomyces fungicidicus* proporcionadas por la presente divulgación proporcionan métodos para producir niveles mejorados de enduracidina. Este avance técnico en el estado del arte permite importantes ahorros en los costos asociados con la producción de enduracidina. En algunos ejemplos, los métodos para producir enduracidina comprenden cultivar una cepa recombinante descrita de *Streptomyces fungicidicus* en condiciones suficientes para producir enduracidina. En algunos ejemplos, el método comprende además aislar la enduracidina desde el medio de cultivo después del cultivo. En algunos ejemplos, el método comprende además determinar la actividad antibacteriana de la enduracidina producida, tal como mediante análisis por HPLC o

bioensayo usando el *S. aureus* ATCC 29213 o *Bacillus subtilis* ATCC 6633 como microorganismos indicadores.

En algunos ejemplos, la enduracidina es producida por una cepa de *Streptomyces fungicidicus* divulgada utilizando las condiciones de fermentación como se describió previamente para la producción de enduracidina (Higashide *et al. J. Antibiot.* 21: 126-137, 1968). Después de la producción, los compuestos se pueden purificar y/o analizar incluyendo el análisis por HPLC como se describe en el Ejemplo 1. Los métodos para producir enduracidina y recolectar este compuesto a partir de un medio de crecimiento se pueden encontrar en la Patente de Estados Unidos 4.465.771, que se incorpora aquí como referencia en su totalidad.

En algunos ejemplos, una cepa de *Streptomyces fungicidicus* descrita se cultiva en caldo de soja tríptico (TSB) en un agitador (como a 225 rpm y 30 °C durante 48 horas) y luego se transfiere a un medio de producción de enduracidina (EPM, Tabla 1 a continuación) durante un período de tiempo para la fermentación continua, por ejemplo, durante al menos cinco días y hasta once días, incluidos 5, 6, 7, 8, 9, 10 u 11 días de fermentación continua. En algunos ejemplos, la producción de enduracidina por las cepas de tipo silvestre y derivadas se realiza en fermentadores automáticos.

Tabla 1. Composición del Medio de Producción de Enduracidina (EPM) (pH 6.7)

Ingrediente	Concentración (%)
Almidon soluble	1.5
Glucosa	1.0
Harina de maíz	2.5
Harina de gluten de maíz	2.0
Licor de maíz fermentado	0.25
Cloruro de sodio	0.25
NaH ₂ PO ₄	1.3
KH ₂ PO ₄	0.05
(NH ₄) ₂ SO ₄	0.15
CaCO ₃	0.5
Lactosa	0.5
ZnCl ₂	0.005
Aceite de pollo	0.7

En algunos ejemplos, la biomasa de *Streptomyces fungicidicus* se produce mediante un proceso de fermentación en fermentadores industriales de diseño sanitario de tanque profundo con sistemas para monitorear y controlar el pH, la temperatura, el oxígeno, la aireación y la agitación. Por ejemplo, cada lote fermentado de *S. fungicidicus* se inicia a partir de un stock de semillas de trabajo caracterizado y controlado de la semilla de producción almacenada en un lugar seguro y mantenida en un ambiente de baja temperatura.

En algunos ejemplos, el proceso de fermentación ocurre en una o más etapas, como las siguientes tres etapas y, opcionalmente, se puede seguir con un procesamiento adicional corriente abajo:

Etapas I:

Los cultivos de semillas de trabajo establecidos caracterizados se utilizan para iniciar un lote de fermentación. Se recuperan de uno a cinco viales de semillas congeladas del almacenamiento a baja temperatura y se descongelan naturalmente o se colocan en un baño de agua entre 28 °C y 32 °C hasta que los contenidos se descongelan. Los cultivos descongelados se transfieren asépticamente a agua estéril mantenida a temperatura ambiente y se mezclan suavemente para volver a suspender el cultivo.

Etapa II:

El cultivo resuspendido se transfiere asépticamente a 0,005 m³-0,05 m³ de medio de siembra. El medio de siembra está compuesto de glucosa (0.1-1.0 g / L), Dextrina (0.1-3 g / L), licor de maíz fermentado (0-5.0 mL / L), polvo de soja (1-5.0 g / L), sulfato de amonio (0.1-0.5 g / L), fosfato monopotásico (0.13-0.54 g / L), sulfato ferroso (0.00-0.5 g / L), hidróxido de potasio (0.13 mL / L), carbonato de calcio (1- 2 g / L), agente desespumante a base de silicona (0,1 ml / L), agua, c. s. El medio se esteriliza a 125 °C-128 °C durante 30-45 minutos y luego se enfría a 28 °C-32 °C. El volumen del medio se ajusta con agua estéril al volumen de trabajo deseado. El pH se ajusta a 6.5-7.0.

Los parámetros operativos del ciclo de ampliación de semillas incluyen: Temperatura de incubación de 28 °C $\pm$ 2 °C, una presión interna de 1.0 $\pm$ 0.5 kg/cm², una tasa de aireación de 3 $\pm$ 2 Nm³/min y una tasa de agitación de aproximadamente 80 rpm, según el tamaño y la configuración del recipiente. El pH, el consumo de oxígeno y la viscosidad se monitorean pero no se controlan. El cultivo se cultiva durante 40-80 horas antes de transferirlo al fermentador de producción principal. La viscosidad en el momento de la transferencia debe oscilar entre 200-600 cps, y el pH debe ser $\leq$ 6.0, y debe haber un aumento en el consumo de oxígeno. El cultivo de semillas se transfiere asépticamente al medio de fermentación principal para completar el ciclo de fermentación.

Etapa III:

La composición del medio fermentador de producción (10 m³-250 m³) incluye componentes naturales y químicos como harina de maíz (13.0-15.0% p/v), harina de gluten de maíz (3.0-6.0% p/v), harina de semilla de algodón (0.1-0.3% p/v), licor de maíz (0,1-0,6% v/v), cloruro de sodio (0,3% v/v), sulfato de amonio (0,25-0,6% p/v), ácido láctico (0-0,5% v/v %), cloruro de zinc (0.01% p/v), sulfato ferroso (0.0-0.02% p/v), hidróxido de potasio (0.20-0.5% v/v), sulfato de calcio (0.0-0.5% p/v), carbonato de calcio (0,5% p/v), amilasa (0,02-0,06% p/v), hidróxido de potasio (0,05% v/v), aceite vegetal (0,5-2,0% v/v), agente desespumante y agua, c. s. Los ingredientes se añaden según el orden indicado. Agregar agua a los ingredientes, luego calentar a 70-90 °C para permitir que la enzima descomponga los carbohidratos complejos durante 15 minutos a temperatura. Agregar los ingredientes restantes, ajustar el pH a 6.6-6.8 y agregar agua c. s., esterilizar a 125 °C-128 °C durante 25-50 minutos para esterilizar el medio. Enfriar los medios a 25 °C-32 °C y agregar agua c. s., volumen de trabajo.

Transfiera los contenidos desde el fermentador de semillas al medio de fermentación principal y ajustar el fermentador a las siguientes condiciones: Temperatura $28\text{ }^{\circ}\text{C} \pm 3\text{ }^{\circ}\text{C}$, tasa de aireación 20-60 Nm³/min, presión interna 0.1-1.0 kg/cm², tasa de agitación equivalente a aproximadamente 1.85 kW/m³. La velocidad de aireación, la presión interna y las velocidades de agitación se ajustan según sea necesario para garantizar que el oxígeno disuelto no sea un límite de velocidad determinado. Controlar cuidadosamente la espuma durante todo el ciclo para evitar la contaminación o la salida. Comenzar a controlar el pH después de que aumenta la demanda de oxígeno. Los siguientes parámetros son controlados y/o monitoreados durante todo el ciclo de fermentación: pH, aeración, oxígeno disuelto, CO₂, viscosidad, pureza, velocidad de agitación, presión interna y azúcar residual. Mantener el pH en 6.8 hasta que cese el crecimiento de las bacterias, luego permitir que el pH cambie naturalmente hasta la cosecha. El ciclo de fermentación típico es de 210-300 horas. El cultivo está listo para ser cosechado cuando la potencia es mayor que 5,000 μL /L, el pH aumenta a 7.5 o más, la viscosidad disminuye y la demanda de oxígeno cesa.

La fermentación se cosecha calentando el cultivo a $70\text{ }^{\circ}\text{C}$ durante 30 minutos para inactivar las bacterias, y luego enfriar los fluidos de cosecha a $25\text{ }^{\circ}\text{C}$ - $32\text{ }^{\circ}\text{C}$.

En algunos ejemplos, el procesamiento posterior incluye eliminar el agua de la biomasa, secar la biomasa y formular la biomasa seca en una premezcla.

Depósitos de Material Biológico

Los siguientes materiales biológicos han sido depositados bajo los términos del Tratado de Budapest con The American Type Culture Collection, y se les han dado los siguientes números de acceso:

<u>Deposito</u>	<u>Número de ingreso</u>	<u>Fecha de depósito</u>
BM38-2-18pfrd	PTA-124007	2 de marzo de 2017
BM38-2-24 / 16	PTA-124006	2 de marzo de 2017

Las cepas anteriores se han depositado en condiciones que aseguran que el acceso al cultivo estará disponible durante la tramitación de esta solicitud de patente a una determinada por el Comisionado de Patentes y Marcas que tendrá derecho a ella de acuerdo con el artículo 37 C.F.R. §1.14 y 35 U.S.C. §122. Los depósitos representan un cultivo sustancialmente puro de las cepas depositadas. Los depósitos están disponibles según lo exigen las leyes de

patentes extranjeras en los países donde se archivan las contrapartes de la solicitud en cuestión o su progenie. Sin embargo, debe entenderse que la disponibilidad de un depósito no constituye una licencia para practicar la invención en cuestión, en derogación de los derechos de patente otorgados por acción gubernamental.

Los siguientes ejemplos no limitantes se proporcionan para ilustrar ciertas características y/o modalidades particulares. No se debe interpretar que estos ejemplos limitan la divulgación a las características o modalidades particulares descritas.

EJEMPLOS

Ejemplo 1

Materiales y Métodos para la Producción Mejorada de Enduracidina

Este ejemplo proporciona métodos representativos para la producción mejorada de enduracidina.

Cepas bacterianas, plásmidos, fosmidos y condiciones de cultivo.

Streptomyces fungicidicus B-5477 (ATCC 21013) y *Escherichia coli* S17-1 (ATCC 47055) se adquirieron de ATCC. Intervet/Merck Animal Health (MAH) proporcionó la cepa BM38-2 de *S. fungicidicus* (ATCC PTA-122342) y los estándares de enduracidinas A y B. Las cepas de *E. coli* DH5 α (Life Technologies, Inc.), EPI300 (Epicenter) y XL10-Gold (Stratagene) se utilizaron como hospedadores para los plásmidos de *E. coli*, fosmids y vectores de lanzadera de *E. coli*-*Streptomyces*. Plásmidos pSET152 (Bierman et al., Gene 116: 43-49, 1992, que se incorpora aquí como referencia en su totalidad) y el pIJ773 proporcionados por el profesor Keith Chater (JIC, Norwich, Reino Unido). El plásmido pWHM860 que alberga el promotor ermE* ρ fue proporcionado por el profesor Bradley Moore (UCSD, San Diego). ISP2 (Medio 2 ISP DifcoTM), ISP4 y TSB (Caldo de Soya Tríptico BactoTM) fueron adquiridos de VWR. Los cebadores utilizados para la PCR y la secuenciación del ADN se sintetizaron a partir de Fisher y Sigma-Aldrich. Los medios y las condiciones de cultivo para cultivar *S. fungicidicus* fueron descritos por Higashide. (*Journal of Antibiotics*, 21: 126-137, 1968). Todos los procedimientos de *E. coli* se realizaron de acuerdo con los protocolos estándar.

Aislamiento y manipulaciones de ADN.

Para preparar ADN genómico a partir de *S. fungicidicus* B-5477, BM38-2 (ATCC PTA-122342) y cepas recombinantes y mutantes derivadas para secuenciación, construcción

de biblioteca de fosmidos, subclonación y PCR, se inocularon y cultivaron esporas recién cosechadas desde cepas individuales. 100 ml de medio líquido TSB suplementado con MgCl_2 5 mM y glicina al 0.5%. El cultivo representativo se realizó en matraces Erlenmeyer de 500 ml en una incubadora rotatoria con agitación a 225 rpm y 30 °C durante 48 a 72 horas. Las células miceliales se recogieron mediante centrifugación a 4000 rpm y 4 °C durante 15 minutos. El sobrenadante se descartó y el sedimento se lavó sucesivamente una vez con 10.3% de sacarosa y dos veces con Tris-HCl 10 mM y etilendiaminotetraacetato de disodio 1 mM (EDTA), pH 8.0 (tampón TE). Las células húmedas, equivalentes al volumen de 80 μL de agua, se distribuyeron en tubos de microcentrífuga estériles de 1.5 ml.. Después de agregar 300 μL de la solución de lisis que contiene 200 μL de Tris-HCl 10 mM y EDTA 1 mM, pH 8.0 y sacarosa 0.3 M (tampón TES), 50 μL de EDTA 0.5 M, 50 μL de lisozima (50 mg / mL), los tubos se incubaron a 37 °C durante 30 a 60 minutos hasta que la solución se volvió viscosa. A continuación, se agregaron a cada tubo 5 μL de proteinasa K (20 mg/mL) y 180 μL de dodecilsulfato de sodio al 10% (SDS). Después de una mezcla suave pero completa, las soluciones se incubaron a 37 °C durante 90 minutos. Luego, se agregaron 80 μL de bromuro de cetil trimetil amonio (CTAB) al 10%. Después de mezclar a fondo, los tubos se incubaron a 65 °C durante 10 minutos. Las soluciones se extrajeron dos veces con 600 μL de fenol/cloroformo/ alcohol isoamílico (25/24/1). El ADN genómico en las fases acuosas superiores se recuperó y precipitó con 0.6 volúmenes de isopropanol. El ADN genómico recogido se lavó dos veces con etanol al 70%. Después de secar a temperatura ambiente durante 10 minutos, el ADN genómico se disolvió en 50 a 100 μL de agua mineral. La alta calidad de la preparación de ADN genómico se confirmó mediante digestión con HindIII y Sau3AI, que mostró una digestión completa y ninguna degradación del ADN genómico no digerido por electroforesis en gel de agarosa al 0.8%. El ADN genómico agrupado se digirió posteriormente con RNasa para eliminar la contaminación por ARN. La pureza y la cantidad del ADN genómico se determinaron con un espectrofotómetro Nanodrop. Se realizaron manipulaciones de ADN de estreptomicetos generales, incluida la electroforesis en gel de agarosa, y se utilizaron los kits QIAprep Spin Miniprep (Qiagen) para preparar plásmidos y fosmids de cepas de E. coli. Las endonucleasas de restricción, la ADN ligasa, la ADN polimerasa, la transposasa, la enzima Klenow, la fosfatasa alcalina y la ligasa se adquirieron de Biolabs, Invitrogen, Epicenter y Roche, y se utilizaron de acuerdo con las recomendaciones del fabricante. Los fragmentos de ADN se purificaron utilizando kits de extracción de gel QIAquick.

PCR.

La PCR de la colonia se realizó de la siguiente manera: se inocularon esporas de colonias candidatas mutantes independientes en cultivo líquido TSB. Después de crecer durante 48 a 72 horas, los micelios se recogieron por centrifugación y se lavaron dos veces con tampón TE (Tris 10 mM, EDTA 1 mM), pH 8.0. Los micelios se resuspendieron en H₂O estéril y se usaron como plantilla en la mezcla de reacción de PCR en un volumen final de 100 µL que contenía 60 µL de micelios, 150 pmol de cada cebador, 20 µL de tampón A rico en AccuPrime GC 5X (Invitrogen), y 1 µL de Polymix (agregado a 80 °C) desde el sistema PCR de plantilla larga Expand (Roche). PCR se realizó de la siguiente manera: 1 ciclo a 95 °C durante 3 minutos, 30 ciclos a 95 °C durante 1 minuto, a 55 °C durante 1 minuto y a 72 °C durante 2 minutos. La reacción se terminó con un ciclo de extensión a 72 °C durante 10 minutos. Los productos de la PCR se purificaron en gel y se secuenciaron. La PCR general se realizó de manera similar a lo descrito anteriormente, excepto que el ADN genómico aislado, el ADN plásmido / fosmid se usó como plantilla en lugar del uso directo del ADN liberado de las colonias de micelios sin purificación previa.

Construcción del plásmido de expresión integrativa pXY152-endorf24

Con el fin de expresar ectópicamente el gen regulador putativo *orf24* del grupo de genes de enduracidina en cepas de *S. fungicidicus* tipo silvestre y BM38-2 (ATCC PTA-122342), *orf24* se clonó en el plásmido integrativo pXY152 derivado de pXY152aR20 (Yin, J. Productos Naturales, 73: 583-589, 2010, que se incorpora en la presente como referencia en su totalidad) *orf24* se amplificó por PCR a partir de ADN genómico de *S. fungicidicus* utilizando el cebador directo (End24Ndpf): 5'-CCACCAC**ATATG**GAAATAAGTTCGCTCTCCA-3' (SEQ ID NO:1, el sitio NdeI está en negrita) y el cebador inverso (End24ERpr:5'-GTGTGT**GAATTC**CCTCGTTCACCCGGCCAGATG-3' (SEQ ID NO: 2, el sitio EcoRI está en negrita). El producto de la PCR se digirió con NdeI y EcoRI. El fragmento *orf24* purificado en gel se ligó luego con el vector pXY152 restringido de manera similar. El plásmido resultante se designó pXY152-endorf24 (FIG. 2, SEC ID NO: 3). Se confirmó que el inserto *orf24* está libre de errores por secuenciación.

Construcción del plásmido pXY300-orf18ifd para la eliminación en marco de *orf18*

pXY300-orf18ifd se construyó mediante la clonación de dos fragmentos que flanquean *orf18* y están destinados a la eliminación en pXY300, un vector de lanzadera

sensible a la temperatura de conjugación de *E. coli*-*Streptomyces* que contiene el gen de resistencia a la tiosstreptona (*tsr*) para la selección en *S. fungicidicus*. Una secuencia de 2 kb “ascendente” y una secuencia de flaqueo “descendente” de 2 kb, y una secuencia “corriente abajo” de 2 kb, designadas *orf18ifdNP* y *orf18ifdPH*, respectivamente, que flanquean *orf18* se generaron mediante PCR utilizando el ADN genómico de *S. fungicidicus* como plantilla y dos conjuntos de cebadores. El fragmento *orf18ifdPH* se amplificó utilizando los cebadores delantero y trasero (*Ifdenorf18pf1*, 5'-TTATTGA**AAGCTT**GCCGGGGCCGACGCGGCGGGCGGCCT-3'

SEQ ID NO: 4), *Ifdendorf18pr1*, 5'-GTTGTT**TAATTA**AAACACCAGGCCTCCTGGGGTG-3'

SEQ ID NO: 5), los sitios *HindIII* y *PacI* están en negrita). El fragmento *orf18ifdNP* se amplificó utilizando los cebadores delantero y trasero (*Ifdendorf18pf2*, 5'

-TTTATAT**TAATTA**ATGACCCTTCCGTCCCGCCCCCGAT-3'

SEQ ID NO: 6), *Ifdendorf18pr2*, 5'

-TTTGGT**GCTAGCT**GGTCGTGGCGCTGTTCC-3'

SEQ ID NO: 7), los sitios *PacI* y *NheI* están en negrita). Estos dos fragmentos de PCR se restringieron de forma apropiada y se ligaron simultáneamente con el vector pXY300 preparado por digestión con *NheI* y *HindIII*, para proporcionar el plásmido pXY300-*orf18ifd* (FIG. 3, SEC ID NO: 8). El inserto de eliminación sin errores en el marco de pXY300-*orf18ifd* se confirmó mediante secuenciación.

Construcción del plásmido pKS-T-*orf18pfrd*-AmR para la eliminación de *orf18* y sus regiones flanqueantes.

El fragmento *oriT* se amplificó por PCR a partir del plásmido pIJ773 utilizando el cebador directo (*Oritnhexbahd3f*, 5'

-AGCACAGCTAGCTTCTAGA**AAGCTT**CATTCAAAGGCCGGCA-3' (SEQ ID NO: 9) el sitio *HindIII* está en negrita) y el cebador inverso (*Oriter1pstxhor*, 5'

-GCCAGTGAATTCTGCAG**CTCGAG**CAGAGCAGGATTCCCGTTGA-3' (SEQ ID NO: 10), el sitio *XhoI* está en negrita). El fragmento *oriT* se digirió con *HindIII* y *XhoI*, se purificó en gel y luego se ligó en el vector derivado pBluescript II KS restringido de manera similar para producir el plásmido pKS-T (Alting-Mees y Short, Nucleic acid Research, 17: 9494, 1989). El inserto del plásmido pXY300-*orf18ifd* se cortó por digestión con *NheI* y *HindIII*, se purificó en gel y luego se ligó con el plásmido pKS-T linealizado con *NheI* y

HindIII para proporcionar el plásmido pKS-T-orf18ifd (FIG. 4, SEC ID NO: 11). Un fragmento de 1 kb que lleva *aac(3)IV*, el gen de resistencia a apramicina (*amR*), se amplificó de pIJ773 utilizando un cebador directo (ApraNcoIpf, 5'-GAATGGCCATGGTTCATGTGCAGCTCCAT-3' (SEQ ID NO: 12), el sitio NcoI está en negrita) y el cebador inverso (ApraBamHIpr, 5'-TCTCGAGGATCCGAATAGGAACTTCGGAAT-3' (SEQ ID NO: 13), el sitio BamHI está en negrita). La digestión del fragmento AmR y el plásmido pKS-T-orf18ifd con NcoI y BamHI prepararon tanto el inserto como el vector para la ligadura. El plásmido resultante se designó pKS-T-orf18pfrd-AmR (FIG. 5, SEC ID NO: 14).

Construcción del plásmido pKS-T-orf18ifd-AmR (NS) para la eliminación en marco de *orf18*.

El inserto de pXY300-orf18ifd se cortó mediante digestión con *NheI* y *HindIII*, se purificó en gel y luego se ligó con el vector linealizado *SpeI* y *HindIII* pBluescript II KS para producir un plásmido pKS-orf18ifd. El fragmento *oriT* se amplificó por PCR utilizando el cebador directo (Oritnhexbahd3f, 5'-AGCACAGCTAGCTTCTAGAAGCTTCATTCAAAGGCCGGCA-3' (SEC ID NO: 15), el sitio HindIII está en negrita) y el cebador inverso (oriTXhNdSpr, 5'-AGGCAGCTCGAGCATATGACTAGTCAGAGCAGGATTCCCGTTGA-3' (SEQ ID NO: 16), los sitios XhoI, NdeI y SpeI están en negrita). El fragmento *oriT* se digirió con *XhoI* y *HindIII*, se purificó en gel y luego se ligó con el plásmido pKS-orf18ifd restringido de manera similar para obtener un plásmido pKS-orf18ifd-T. Un fragmento de 1 kb que lleva el gen *aac(3)IV* que confiere resistencia a la apramicina (AmR) se amplificó a partir de pIJ773 por PCR usando el cebador directo (ApraNdepf, 5'-GAATGGCATATGGTTCATGTGCAGCTCCAT-3' (SEQ ID NO: 17), el sitio NdeI está en negrita) y el cebador inverso (ApraSpeIpr, 5'-TCTAGAACTAGTGAATAGGAACTTCGGAAT-3' (SEC ID NO: 18), el sitio SpeI está en negrita). El plásmido pKS-orf18ifd-T se linealizó mediante digestión con *NdeI* y *SpeI* y luego se ligó con el fragmento AmR restringido de manera similar para generar el plásmido pKS-orf18ifd-T-AmR(NS) (FIG. 6, SEC ID NO: 19).

Conjugación intergenérica, procedimientos de alteración génica basados en pXY300 y basados en pKS.

Los plásmidos de alteración génica se introdujeron individualmente en *E. coli* S17-1 por transformación y luego se transfirieron a *S. fungicidicus* o sus derivados mediante conjugación. Brevemente, las esporas de *S. fungicidicus* recién recolectadas fueron pregerminadas y las células de *E. coli* S17-1 se cultivaron durante la noche a 37° C en un caldo Terrific. Se realizaron diluciones en serie de la suspensión de esporas germinadas y se mezclaron 100 ml de cada dilución con un volumen igual de *E. coli* S17-1 que albergaba los plásmidos de interrogación basados en pXY300. Las soluciones se colocaron en placas de agar ISP4 con adición de 10 mM de MgCl₂ y se incubaron durante 22 horas bien sea a 30 o 37 °C. Cada placa se cubrió con 3 ml de agar nutriente blando que contenía nalidixato de sodio y apramicina (0,5 mg/ml) y se incubó adicionalmente a 30 °C durante aproximadamente una semana. Los exconjugantes aislados que sobrevivieron se purificaron por sembrado en placas de agar ISP4 suplementadas con nalidixato de sodio y apramicina (50 µg/ml cada una).

Para realizar los estudios de alteración génica con plásmidos basados en pXY300, los exconjugantes se cultivaron primero en medio líquido TSB que contenía apramicina (5 µg/mL) a 30 °C durante 24 horas, momento en el cual los micelios se recolectaron, homogeneizaron y se usaron para inocular el medio líquido TSB suplementado con apramicina (5 µg/mL). Después de 3-6 días de incubación a 40 °C, los micelios se homogeneizaron y se colocaron en placas de agar ISP4 que contenían apramicina (50 µg mL) y se incubaron a 30 °C durante una semana. El ADN genómico se aisló de colonias sobrevivientes individuales seleccionadas al azar y se analizó mediante bien sea PCR o Southern blot para confirmar que se había producido una alteración de cruce simple o doble. Para la ruptura de genes basada en pKS y los plásmidos de eliminación en el marco, los exconjugados se pasaron a través de tres rondas sucesivas de incubaciones en placas de agar ISP4 para la esporulación sin adición de ninguna selección de antibióticos para estimular la conversión a recombinantes de doble cruce. Los exconjugantes basados en pKS no se pasaron a través de la selección de temperatura a 40 °C. La construcción correcta de todos los mutantes se confirmó mediante PCR y/o análisis de transferencia Southern.

Construcción de los plásmidos de expresión integrativa pXY152-endorf24-camtsr y pXY152-endorf24-blatsr.

Para expresar ectópicamente *orf24* en el mutante resistente a apramicina que lleva la delección de *orf18* y sus regiones flanqueantes, se diseñó el plásmido de expresión integradora pXY152-endorf24-blatsr. Para construir este plásmido, un casete (camtsr) que alberga el gen de resistencia al cloranfenicol y el gen de resistencia a la tiostrepton (tsr) se escindió de un plásmido pUC57 derivado por la digestión con *SacI* y *NheI*. El casete de camtsr se ligó luego con el plásmido pXY152-endorf24 linealizado con *SacI* y *NheI* para proporcionar un nuevo constructo pXY152-endorf24-camtsr (la FIG. 7, SEC ID NO: 20). Un gen de resistencia a la ampicilina (*bla*) se amplificó por PCR a partir de pBluescript KS utilizando el cebador directo (amp2956SwaIpf, 5'-GTGGCA**ATTTAAAT**GGAATGTGCGCGGAA-3' (SEQ ID NO: 21), el sitio *SwaI* está en negrita) y el cebador inverso (amp1973SacIpr, 5'-TATATAG**AGCTCA**ACTTGGTCTGACAGTTAC-3' (SEQ ID NO: 22), el sitio *SacI* está en negrita). *bla* se clonó luego en los sitios *SacI* y *SwaI* de pXY152-endorf24-camtsr reemplazar el cassette camtsr con blatsr. El plásmido de expresión conyugal resultante se designó pXY152-endorf24-blatsr (FIG. 8; SEC ID NO: 23).

Construcción del cassette Tn5AT para la mutación de transposones in vitro

El cassette Tn5AT fue diseñado para combinar tres elementos genéticos: el transposón *Tn5*, *oriT* y *aac3 (IV)*. La *Tn5* es reconocida de manera específica y única por la transposasa *Tn5* (Epicentro) y se inserta fácilmente en el ADN de *Streptomyces* con alto contenido de G + C clonado en plásmidos y fosmids de *E. coli* (también referido en la Patente de EE.UU. N° 8.188.245, que se incorporó a la presente como referencia). *oriT* es necesario para la transferencia conyugal de ADN de *E. coli* S17-1 a *Streptomyces* y *aac3(IV)* es un marcador de selección bifuncional de *E. coli-Streptomyces* que confiere resistencia a apramicina. Tanto *oriT* como *aac3(IV)* se escindieron del plásmido pIJ773 como un fragmento *XbaI* y luego se clonaron en el plásmido pMODTM-2 (MCS) (Epicentro) del transposón, previamente linealizado con *XbaI*. Los plásmidos resultantes pXYTn5ATa y pXYTn5ATb solo difieren por la orientación del fragmento *XbaI* y se utilizaron para preparar el cassette de Tn5AT mediante digestión con *PvuII* según las especificaciones del fabricante.

Mutación y selección de transposones in vitro de fosmids pXYF24D3 y pXYF148D12 mutagenizados

Para generar una biblioteca de fosmids mutagenizados al azar que llevan segmentos del grupo de biosíntesis de enduracidina para estudios de reemplazo de genes, se realizaron estudios de mutación por inserción de transposones in vitro de fosmids pXYF24 y pXYF148. Dos genes reguladores putativos de la biosíntesis de enduracidina, *orf18* y *orf24*, residen en los insertos de fosmids pXYF24 y pXYF148, respectivamente (GenBank número de acceso DQ403252). La reacción del transposón in vitro se realizó a 37 °C durante 2 horas después de mezclar 10 µL (0,5 µg) de ADN de plantilla de fosmid, 2 µL (20 ng) de ADN de cassette Tn5AT, 2 µL de 10 x tampón de reacción, 1 µL de transposasa de Tn5 y 5 µL de agua estéril. La transformación de las células competentes de *E. coli* EPI300TM-T1^R (Epicentro) con la mezcla de reacción de transposón se realizó por electroporación. Los fosmids mutagenizados fueron seleccionados en placas de agar LB suplementadas con 100 µg/mL de apramicina. Las placas se incubaron durante la noche a 37 °C y las colonias supervivientes se seleccionaron al azar y se cultivaron en un cultivo líquido LB con la adición de 100 µg/mL de apramicina. El ADN de fosmid mutagenizado de estas colonias y el fosmid pXYF24 o pXYF148 de control se digirieron con *HindIII* y se analizaron por electroforesis en geles de agarosa al 1%. El cassette de Tn5AT contiene un único sitio *HindIII* que es útil cuando se analizan eventos de alteración únicos contra múltiples sobre el inserto de fosmid. No existen sitios *HindIII* presentes en los insertos de fosmid de pXYF24 o pXYF148, y solo un sitio *HindIII* está presente en el vector del fosmid. Por lo tanto, la digestión con *HindIII* identifica fácilmente fosmids con una inserción única de Tn5AT por la presencia de dos bandas en el gel. Las colonias que portaban fosmidos mutagenizados con una inserción de transposón única se seleccionaron al azar y se hicieron crecer en un cultivo líquido de LB para permitir el aislamiento de fosmid y la identificación del gen alterado. La selección se realizó mediante análisis de secuencia utilizando el cebador 5'- AAGGAGAAGAGCCTTCAGAAGGAA-3' (SEC ID NO: 24), que corresponde a una región del gen de resistencia a la apramicina. De esta manera, se encontró que el fosmid pXYF24D3 y pXYF148D12 tienen Tn5AT insertado en *orf18* en la posición de nucleótido 26386 y *orf24* en la posición de nucleótido 34333 (número de acceso de GenBank DQ403252), respectivamente.

Interrupción de inserción de *orf18* y *orf24* en el *S. fungicidicus* tipo silvestre ATCC 21013.

Los fosmids de sustitución de genes pXYF24D3 y pXYF148D12 se transformaron por separado en *E. coli* S17-1 por electroporación y luego se introdujeron en *S. fungicidicus* mediante conjugación intergenerica (Mazodier, J. Bacteriology 171: 3583-3585, 1989, que se incorpora aquí como referencia en su totalidad). Las colonias exconjugantes que sobrevivieron a la selección de apramicina se pasaron a través de tres rondas sucesivas de esporulación sin selección de antibióticos en placas de agar ISP2 para crear la cepa mutante estable a través de recombinación homóloga de doble cruce. Las esporas resultantes se agruparon, se diluyeron y se sembraron en placas de agar ISP2 suplementadas con 50 µg/mL de apramicina para confirmar la resistencia a la apramicina y para el uso en cultivos de semillas y fermentación para la producción de enduracidina. La cepa mutante con la alteración de inserción de *orf18* en *S. fungicidicus* tipo silvestre se designó como SfpXYF24D3 y la cepa mutante con la alteración de inserción de *orf24* en *S. fungicidicus* tipo silvestre se designó como SfpXYF148D12.

Producción de enduracidina a escala de laboratorio y en fermentador de 10 litros.

Las condiciones de fermentación en matraz de agitación de laboratorio para la producción de enduracidina en *S. fungicidicus* tipo silvestre, BM38-2 (ATCC PTA-122342) y cepas derivadas fueron las descritas por Higashide, (J. Antibióticos, 21: 126-137, 1968), excepto por los medios de producción de enduracidina que se describieron en una patente (Patente de Estados Unidos N° 4465771). Para la fermentación a escala de laboratorio, se utilizaron 5 mL de TSB para la inoculación del cultivo de semilla con esporas de estreptomicetos recién recolectadas. Normalmente, de 5 a 10 mL del cultivo de siembra incubado en un agitador rotatorio a 225 rpm y 30 °C durante 48 horas y luego se transfirió a un medio de producción de enduracidina de 50 mL durante 10 días de fermentación continua. La producción de enduracidina por cepas tipo silvestre y derivadas en condiciones estrechamente controladas también se llevó a cabo en fermentadores automáticos de 10 litros.

Tabla 2: Comparación de los rendimientos de enduracidina (enramicina) en cepas tipo silvestre, mutantes y genéticamente modificadas de *Streptomyces fungicidicus*

<i>Cepa S. fungicidicus</i>	Condiciones de Fermentación	Rendimiento (HPLC)
Tipo silvestre (ATCC21013)	Matraz de Agitación	5-30 mg/L
BM38-2	Matraz de Agitación	60-90 mg/L
SfpXY52endorf24	Matraz de Agitación	60 mg/L
SfpXYF24D3	Matraz de Agitación	40 mg/L
BM38-2.orf18pfrd-AmR	Matraz de Agitación	67 mg/L
BM38-2.24/16	Matraz de Agitación	30-130 mg/L
BM38-2	Fermentador de 10L	80-145 mg/L
BM38-2.24/16	Fermentador de 10L	375 mg/L

Extracción de enduracidina de productos de fermentación para el análisis por HPLC.

Para extraer los metabolitos para el análisis por HPLC de la producción de enduracidina, el micelio fresco se recogió por centrifugación y se lavó con agua desionizada y se resuspendió en un volumen de 5x (relación del metanol acuoso (ml) al peso del micelio húmedo (g)) metanol acuoso al 70% (el pH se ajustó a 3,5 con HCl 1 N). La suspensión se agitó a 200 rpm a temperatura ambiente durante la noche y luego se centrifugó a 4000 rpm a 4 °C durante 20 minutos. Luego, se transfirieron 1,4 mL de sobrenadante de cada muestra a tubos de microcentrífuga individuales de 1,5 mL y se centrifugaron a 13.000 rpm a temperatura ambiente durante 10 minutos. El filtrado se pasó a través de un filtro de jeringa de 0,22 µm y luego se analizó por HPLC. La extracción de metabolitos de micelios producidos en fermentadores de 10L se realizó a pequeña escala, equivalente a las fermentaciones de laboratorio.

Análisis por HPLC y determinación del rendimiento de enduracidina.

Se inyectó una muestra de HPLC de 50 µL preparada como se describe anteriormente en una columna Gemini C₁₈ (5 µm, 4.6x150 mm, Phenomenex, Torrance, CA) unida a una HPLC Shimadzu. La separación se logró mediante un gradiente lineal paso a paso de 18 min con disolvente A: agua + TFA al 0,1% y disolvente B: acetonitrilo. La tasa de flujo fue de 1 mL/ minuto, comenzando con B al 10%, aumentando a B al 40% durante 10 minutos, y luego aumentando de nuevo a B al 95% durante 8 minutos. La región UV de 200 a 300 nm se exploró con un detector de matriz de fotodiodos SPD M20A. Los

rendimientos de enduracidinas se calcularon por comparación con una curva estándar construida a partir de una solución madre de patrones de enduracidina en metanol al 70%. Se utilizó una serie de inyecciones que incluían 2, 4, 6, 8, 10 y 12 µg de enduracidina para construir la curva estándar utilizando la suma de las áreas de absorbancia para las enduracidinas A y B a 230 nm. Se generó una ecuación de regresión a partir de la curva estándar y se usó para calcular los rendimientos de enduracidina.

Evaluación de la actividad antibacteriana.

Se usó *Staphylococcus aureus* (ATCC 29213) como microorganismo indicador en el bioensayo. Las células se utilizaron para inocular el caldo LB, se hicieron crecer a 37 °C durante la noche y luego se mezclaron 100 µL del cultivo con 5 mL del agar superior (mezcla de volúmenes iguales de agar nutritivo y caldo nutriente). El agar superior se superpuso sobre una placa de agar nutritivo en la que se hicieron pocillos adecuadamente espaciados cortando los tapones de agar. Los estándares de enduracidina y las alícuotas de las extracciones de cultivo se disolvieron o diluyeron en MeOH al 50% a una concentración de 20 µg/ml, y se cargaron 100 µL de cada solución en los pocillos. Después de incubar las placas a 37 °C durante 16 horas, se observaron y compararon las zonas de inhibición, y las placas se fotografiaron o almacenaron a 4 °C.

Ejemplo 2

Alteración de *orf18* y *orf24* en *S. fungicidicus* tipo silvestre y efecto en la producción de enduracidina

Este ejemplo describe la alteración de *orf18* y *orf24* en *S. fungicidicus* tipo silvestre y el efecto sobre la producción de enduracidina.

Una secuencia de ADN de 116.000 pb del *S. fungicidicus* ATCC 21013 tipo silvestre que alberga el grupo de genes biosintéticos de enduracidina y sus regiones flanqueantes (EE. UU. Pat. No. 8.188.245 que se incorpora aquí como referencia en su totalidad) se identificó previamente y está disponible en GenBank (número de acceso DQ403252). Entre los 48 marcos de lectura abierta anotados hay ocho genes reguladores putativos: *orf5*, *orf12*, *orf18*, *orf22*, *orf24*, *orf41*, *orf42* y *orf43*. Para descifrar el papel de cada uno de los productos genéticos en la producción de enduracidina, los insertos del fosmid que llevan segmentos del grupo de enduracidina que albergan estos genes reguladores putativos se mutaron al azar

utilizando una inserción mediada por transposones de un marcador de resistencia a apramicina como se describe en el Ejemplo 1.

La selección subsiguiente de la resistencia a apramicina y la ubicación del inserto entre las colonias de *E. coli* que portan fosmidos mutagenizados identificaron pXYF24D3 para transportar el *orf18* alterado y pXYF148D12 para albergar el *orf24* alterado. Una sola mutación de inserción en cada uno de estos fosmids y el sitio de la inserción se confirmó por secuenciación. Estos dos fosmids mutagenizados se introdujeron individualmente por conjugación en la cepa de *S. fungicidicus* tipo silvestre. Los exconjugantes que muestran resistencia a la apramicina se pasaron luego a través de tres rondas de esporulación en agar ISP2 sin la adición de ninguna selección de antibiótico para promover la conversión de la recombinación homóloga de un solo cruce a la mutación de doble cruce. Las cepas mutantes estables resultantes SfpXYF24D3 y SfpXYF148D12 se fermentaron en medio de producción de enduracidina (EPM) a escala de laboratorio en matraces de agitación. El análisis por HPLC de la extracción de metanol al 70% de los micelios a partir de 10 días de fermentación reveló un aumento de 1.3 veces en el rendimiento de enduracidina por la cepa SfpXYF24D3 de *orf18* alterado y la pérdida completa de la producción de enduracidina por la cepa SfpXYF148D12 con un *orf24* alterado. Los extractos de micelios también se evaluaron para la actividad hacia *S. aureus*. La alteración de *orf18* SfpXYF24D3 retuvo la actividad, mientras que la alteración de *orf24* SfpXYF148D12 perdió la actividad hacia *S. aureus*.

Ejemplo 3

Construcción de la cepa recombinante SfpXY152-endorf24 y efecto en la producción de enduracidina

Este ejemplo describe la construcción de la cepa recombinante SfpXY152-endorf24 y la capacidad de esta cepa para producir enduracidina.

La pérdida de producción de enduracidina en la cepa mutante SfpXYF148D12 indicó un posible papel regulador del *orf24*. Una búsqueda BLAST de la base de datos GenBank utilizando la secuencia de proteína de *Orf24* reveló una alta similitud de secuencia con una proteína reguladora específica a vía, StrR, involucrada en la biosíntesis de estreptomicina. Una alineación de secuencias entre *Orf24* y StrR mostró que las proteínas comparten una similitud significativa (54% de identidad de aa, FIG. 9). La pérdida de

producción de enduracidina tras la alteración de *orf24* y la similitud con StrR indican que *Orf24* puede actuar como un activador específico a *vía* en la producción de enduracidina.

Para explorar el papel de *orf24* como una diana reguladora positiva para la mejora de la cepa, el plásmido de expresión integrativa pXY152-endorf24 (FIG. 2) fue construido (Ejemplo 1). Se introdujo el plásmido pXY152-endorf24 en *S. fungicidicus* tipo silvestre mediante conjugación y se examinaron los exconjugantes para determinar el fenotipo de resistencia a apramicina, lo que llevó a la identificación de la nueva cepa recombinante SfpXY152-endorf24. Al menos diez colonias exconjugantes independientes de esta cepa se seleccionaron y purificaron al azar. Estas cepas de colonias llevan el plásmido pXY152-endorf24 integrado en un sitio attB en el cromosoma de *S. fungicidicus* mediante recombinación homóloga de un solo cruce con el sitio attP en el plásmido.

Para investigar los metabolitos producidos por las cepas recombinantes, las esporas de dos cepas de colonias se inocularon en el cultivo de semillas TSB y luego se transfirieron a un medio de producción de enduracidina para la fermentación a escala de laboratorio. El análisis de HPLC de los extractos de metanol al 70% de los micelios recolectados reveló un aumento de 2 veces (60 mg/L) en la producción de enduracidina por ambas cepas recombinantes en comparación con la cepa de tipo silvestre (30 mg/L). Los rendimientos elevados de enduracidina observados en estas cepas de colonias que son capaces de sobreexpresar *orf24* son una evidencia adicional del papel regulador positivo que este gen tiene en la producción de enduracidina y los resultados son consistentes con los obtenidos a partir de la interrupción de *orf24* que condujo a la pérdida de la producción de enduracidina.

Ejemplo 4

Construcción de la cepa BM38-2.24/16 que sobreexpresa *orf24* en *S. fungicidicus* BM38-2 (ATCC PTA-122342) y efecto sobre la producción de enduracidina

Este ejemplo describe la construcción de la cepa BM38-2.24/16 (Número de depósito ATCC PTA-124006), la sobreexpresión de *orf24* en *S. fungicidicus* BM38-2 (ATCC PTA-122342) y el efecto sobre la producción de enduracidina.

Para explorar más a fondo el papel regulador positivo de *Orf24*, se incorporó el plásmido pXY152-endorf24 en el cromosoma de la cepa de producción comercial *S. fungicidicus* BM38-2 (ATCC PTA-122342), como se describió anteriormente para el organismo de tipo silvestre. La selección de exconjugantes que exhiben el fenotipo de

resistencia a apramicina proporciono varias cepas de colonias recombinantes, incluyendo *S. fungicidicus* BM38-2.24 / 16, capaz de producir niveles elevados de enduracidina de hasta 200 mg/L (para un aumento de 3.3 veces sobre BM38-2 (ATCC PTA-122342)) en cultivos de matraz de agitación de laboratorio. *S. fungicidicus* BM38-2-24/16 se seleccionó para una evaluación adicional de la capacidad de producción de enduracidina en base a los rendimientos durante la selección preliminar.

La producción de enduracidina mediante la cepa recombinante de *S. fungicidicus* BM38-2.24/16 en cultivos en matraz de agitación de laboratorio mostró un claro potencial de mejora significativa sobre BM38-2 (ATCC PTA-122342) y también se observó que los rendimientos varían mucho. Para controlar más de cerca las condiciones de cultivo durante el período de crecimiento de 10 días, incluido el pH y el oxígeno disuelto que no se manejan fácilmente en matraces de agitación, la producción se evaluó a través de múltiples corridas en fermentadores de 10L. Bajo estas condiciones más estrechamente controladas, los rendimientos fueron más consistentes y las fermentaciones triplicadas de 10 litros promediaron 375 mg/ml (4.6 veces de BM38-2 (ATCC PTA-122342)). El aumento de los rendimientos de enduracidina en la cepa recombinante de *S. fungicidicus* BM38-2.24/16 (Número de depósito de ATCC PTA-124006) apoya además un papel de aumento de Orf24 en la producción de enduracidina.

Ejemplo 5

Construcción de la cepa mutante de eliminación BM38-1.18pfrd-AmR y el efecto sobre la producción de enduracidina

Este ejemplo describe la construcción de la cepa mutante de eliminación BM38-2.18pfrd-AmR (Nº de depósito ATCC PTA-124007) y el efecto sobre la producción de enduracidina.

orf18 se encuentra en la región corriente arriba del grupo de genes biosintéticos de enduracidina (número de acceso GenBank DQ403252). *Orf18* parece tener un papel negativo en la producción de enduracidina en la medida en que la alteración de inserción del gen en la cepa mutante SfpXYF24D3 elevó el rendimiento de enduracidina. Basándose en esta observación, se diseñaron constructos para la eliminación de *orf18* solo y *orf18* y porciones de sus regiones flanqueantes. Para este propósito, se construyó el plásmido pKS-T-*orf18pfrd*-AmR (FIG. 5). Este plásmido derivado de vector pKS no posee un replicón de

estreptomicetos ni un elemento para la integración en el cromosoma de estreptomicetos. Solo puede intercambiar su inserción con un segmento definido de ADN en el cromosoma del huésped mediante recombinación homóloga de doble cruce. El mapa de inserción de este plásmido se muestra en la FIG. 10. *orf18* y sus regiones flanqueantes que contienen la totalidad del *orf19* y la región que codifica la porción de extremo N-terminal de *orf17* se eliminan en el plásmido pKS-T-*orf18pfrd*-AmR. El brazo izquierdo de 1-kb contiene la región que codifica la porción de extremo C-terminal de *orf17* y su región corriente abajo y el brazo derecho de 1 kb contiene un segmento parcial de *orf20* que codifica la región de extremo N-terminal. Por lo tanto, la eliminación después de una recombinación homóloga de doble cruce dio como resultado una cepa recombinante donde se eliminó la totalidad de *orf18* más *orf19* y la región que codifica la porción de extremo N-terminal de *orf17* y se reemplazó con el gen resistente a la apramicina.

El plásmido pKS-T-*orf18pfrd*-AmR se introdujo de forma conjugada en *S. fungicidicus* BM38-2 (ATCC PTA-122342) y se promovió la recombinación homóloga de cruce simple y doble en placas de agar ISP4 sin suplementación con apramicina. Los exconjugantes que pudieron sobrevivir a la selección posterior de apramicina se purificaron y esta nueva cepa recombinante se designó BM38-2.18pfrd-AmR (No. de Depósito ATCC PTA-124007). Las esporas de esta cepa se inocularon en un medio TSB para el cultivo de semillas y luego se transfirieron a un medio de producción de enduracidina. Después de 10 días de fermentación, los micelios se recolectaron, se procesaron y se analizaron por HPLC. En relación con la cepa precursora BM38-2 (ATCC PTA-122342), se observó un aumento de 1,2 veces en la producción de enduracidina a partir de estas fermentaciones a escala de laboratorio. El aumento relativo en el rendimiento es similar al observado con la cepa SfpXYF24D3 derivada de tipo silvestre y los resultados implican que *orf19* y *orf17*, que flanquean *orf18* y se vieron afectados en la construcción de BM38-2.18pfrd-AmR, tienen poco o ningún efecto sobre la producción de enduracidina. Por lo tanto, el aumento de la producción de enduracidina en la cepa recombinante BM38-2.18pfrd-AmR se debe a la eliminación del papel regulador negativo de *Orf18*.

Con respecto a la eliminación de *orf18* solo con el plásmido pXY300-*orf18ifd* en BM38-2 (ATCC PTA-122342), se encontraron dificultades al seleccionar positivamente los exconjugantes y mutantes simples/dobles con marcador de resistencia a la tiostreptona. Por lo tanto, se usó el vector alternativo pBluescript KS II para construir los plásmidos de administración de reemplazo de genes sin marcadores, tales como pKS-T-*orf18ifd* (FIG. 4)

o pKS-orf18ifd-T, pKS-orf18ifd-T-AmR (NS) (el gen de resistencia a la apramicina se transporta en el vector en lugar de la inserción en *orf18*, véase la FIG. 6).

Ejemplo 6

Desarrollo del vector de inactivación del gen derivado de pKS serie pKS-T-orf18pfrd-AmR

Este ejemplo describe el desarrollo del vector de inactivación del gen derivado de pKS serie pKS-T-orf18pfrd-AmR.

Se desarrollaron una serie de vectores de inactivación de genes derivados de pKS (FIGS. 4, 5 y 6) que poseen la función conjugativa y no requieren pasar transformantes a través de una selección a alta temperatura para eliminar el plásmido como lo requieren algunos otros vectores de alteración de genes. Estos vectores derivados de pKS llevan un replicón no estreptomiceto que permite la replicación en *E. coli* y pueden mantener y seleccionarse con el marcador de resistencia a apramicina en *Streptomyces* y *E. coli* o ampicilina en *E. coli*. Produjeron copiosas copias estables de plásmidos recombinantes en *E. coli* para la conjugación y se han diseñado con varios sitios de restricción únicos y raros encontrados en el ADN de estreptomicetos, tales como PacI, HindIII, NheI y XbaI, que se pueden usar convenientemente para ensamblar el ADN diana al plásmido para la alteración de genes de inserción y los estudios de eliminación en marco.

Ejemplo 7

Desarrollo de vectores de expresión de genes integradores derivados de pSET152 pXY152-endorf24-camtsr y pXY152-endorf24-blatsr

Este ejemplo describe el desarrollo de los vectores de expresión de genes integradores derivados de pSET152 pXY152-endorf24-camtsr (SEQ ID NO: 20) y pXY152-endorf24-blatsr (SEC ID NO: 23).

Se desarrollaron dos vectores nuevos, pXY152-endorf24-camtsr (FIG. 7) y pXY152-endorf24-blatsr (FIG. 8). Poseen funciones conjugativas e integradoras como el vector pSET152, el vector integrador más utilizado para la expresión y complementación de genes de estreptomicetos. Ambos vectores tienen varios sitios de restricción que son raros en el ADN de *Streptomyces* para la clonación y el ensamblaje convenientes del constructo de

expresión. El vector pXY152-endorf24-camtsr se puede mantener y seleccionar en *E. coli* con cloranfenicol a 12,5 µg/mL y en *Streptomyces* con tiostrepton a 50 µg/mL. El vector pXY152-endorf24-blatsr se puede mantener y seleccionar en *E. coli* con ampicilina y en *Streptomyces* con tiostrepton.

Resumen de los Ejemplos 1-6:

Manipulación genética de los genes reguladores de *Streptomyces* y biosíntesis para mejorar la cepa.

Entre los numerosos productores microbianos de productos naturales, aproximadamente el 75% de los antibióticos microbianos conocidos son producidos por actinomicetos. *Streptomyces*, bacterias filamentosas grampositivas, son miembros de la familia de los actinomicetos y son conocidos por su capacidad incomparable para producir una variedad versátil de metabolitos secundarios estructuralmente diversos, farmacológicamente y biológicamente activos. Los policétidos producidos por policétido sintasas (PKS) y productos peptídicos naturales hechos por péptidos sintetasas no ribosomales (NRPS) son representativos.

La investigación sobre la biosíntesis de antibióticos de productos naturales tiene algunos desafíos comunes: primero, cómo superar la baja producción típica de los compuestos precursores o modificados estructuralmente producidos por las cepas de tipo silvestre o modificadas genéticamente; en segundo lugar, cómo activar las numerosas vías biosintéticas del metabolito secundario críptico o huérfano identificadas a partir de secuencias genómicas para poder estudiar la función biológica de los productos. Los avances en el estudio de la biosíntesis de antibióticos de productos naturales en las últimas décadas han indicado que la producción de metabolitos secundarios está regulada por muchas vías. Por ejemplo, los genes biosintéticos precursores y estructuralmente ensamblados (como PKS y NRPS), genes reguladores y genes de auto-resistencia se pueden agrupar en el cromosoma bacteriano. La producción de antibióticos puede estar regulada por genes reguladores específicos de la vía, incluidos activadores y/o represores, genes reguladores ectópicos pleiotrópicos y sistemas reguladores de dos componentes. Las mutaciones que ocurren en cualquiera de estos genes o sistemas reguladores pueden aumentar, disminuir o abolir completamente la producción de antibióticos. La vía biosintética críptica puede activarse por una mutación imprevista que conduce a la producción de un producto previamente desconocido.

La mejora de las cepas puede desempeñar un papel importante en la producción a escala industrial rentable de antibióticos u otros metabolitos microbianos secundarios. Las cepas mutantes capaces de producir rendimientos mayores de metabolitos particulares pueden generarse a través de mutaciones aleatorias, o mediante la ruptura dirigida de genes específicos o mediante la introducción de genes que eliminan los cuellos de botella en una ruta de biosíntesis. La manipulación genética de genes reguladores positivos y negativos, así como los genes biosintéticos, para generar la hiperproducción de metabolitos secundarios específicos ha demostrado ser una estrategia poderosa y altamente exitosa para la mejora de la cepa de actinomicetos.

En la divulgación actual, se demostró el papel regulador positivo de *orf24* y el papel regulador negativo de *orf18* en la producción de enduracidina. La inactivación por inserción específica de *orf24* dio lugar a una pérdida completa de la producción de enduracidina en la cepa recombinante SfpXYF148D12. La sobreexpresión posterior de *orf24* bajo el control del promotor constitutivo fuerte ermE*p en las cepas recombinantes SfpXY152-endorf24 y BM38-2.24/16 condujo a aumentos en los rendimientos de enduracidina de 2 a 4,6 veces aproximadamente. La eliminación de *orf18* y sus regiones flanqueantes, incluyendo la totalidad de *orf19* y una porción de *orf17*, incrementó los rendimientos de enduracidina en 1,2 veces. Estos resultados proporcionaron evidencia genética sólida en apoyo del rol de *orf24* y *orf18* como activador positivo y represor negativo, respectivamente, en la biosíntesis de enduracidina.

Los ortólogos de Orf24 se han confirmado funcionalmente a partir de otras vías de biosíntesis de antibióticos

Una consulta BLAST con la secuencia de proteína de Orf24 contra la base de datos GenBank reveló cientos de visitas (número de acceso de GenBank DQ403252). Muchos muestran una similitud de aminoácidos muy alta (de 60% a 99% de identidad) y se anotan como reguladores de transcripción en la biosíntesis del antibiótico aminoglucósido estreptomicina. Sin embargo, ninguno de este grupo de genes ha tenido la función experimentalmente verificada. El análisis de los resultados de BLAST identificó varias proteínas relacionadas que comparten una menor similitud (más del 40% pero menos del 60% de identidad de aa) con *Orf24* que se caracterizaron funcionalmente. Estos incluyen la proteína StrR bien caracterizada que comparte una similitud menor pero significativa (54% de identidades de aa en 311 aa superpuestas) con Orf24. Se ha demostrado genéticamente y bioquímicamente que StrR funciona como un activador positivo específico de la vía de la

expresión de los genes de biosíntesis de estreptomicina en *Streptomyces griseus*. StrR representa una familia de activadores específicos de la vía, algunos de los cuales se han caracterizado por bien sea manipulación genética o estudios bioquímicos. La FIG. 11 muestra la alineación de Orf24 con seis proteínas de tipo StrR de actinomicetos funcionalmente confirmadas. Un dominio de unión al ADN típico y altamente conservado de hélice-giro-hélice (HTH) está presente en las siete proteínas como se subraya en la FIG. 11. Orf24 también comparte una similitud de secuencia significativa (54% de identidades aa) con Tei15*, un activador específico de la vía que rige la biosíntesis del antibiótico glicopeptídico no tribosómico teicoplanina. Tei15* regula positivamente la transcripción de al menos 17 genes en el grupo de teicoplanina. El *Actinoplanes teichomyceticus* de tipo silvestre produce 100 mg/L de teicoplanina aproximadamente, mientras que las cepas recombinantes genéticas, derivadas del *A. teichomyceticus* precursor y que contienen *tei15** expresado bajo el control de diferentes promotores, aumentaron el rendimiento de teicoplanina a 1 g/L en el caso del promotor ermE*p y a 4 g/L en el caso del promotor del gen de resistencia a apramicina nativo.

Como se ilustra en la FIG. 11, Orf24 también comparte una similitud de secuencia significativa (54% de identidad aa) con Bbr, del grupo de biosíntesis de antibióticos de balhimicina glucopéptido; a KasT (identidades del 50% aa) que rigen la expresión de los genes de biosíntesis de kasugamicina del antibiótico aminoglucósido; y NovG (45% de identidades aa), el activador específico de la vía involucrado en la biosíntesis de novobiocina. El mutante Δ novG produjo solo un 2% tanto de novobiocina como de tipo silvestre y la sobreexpresión de *novG* de un plásmido de copias múltiples en la cepa recombinante llevó a un aumento de tres veces en la producción de novobiocina. Orf24 también comparte un 42% de identidad aa con SgcR1, uno de los cuatro genes reguladores (*sgcR1*, *sgcR2*, *sgcR3* y *sgcR*) confirmados experimentalmente que participan en la producción del antibiótico antitumoral C-1027 en *S. globisporus*. La sobreexpresión de *sgcR1* en *S. globisporus* SB1022 aumentó el rendimiento de C-1027 aproximadamente siete veces en comparación con la cepa de tipo silvestre. La sobreexpresión del regulador positivo *sgcR3* en una cepa recombinante resultó en un aumento del 30-40% en la producción de C-1027. En contraste, la inactivación del regulador negativo *sgcR* condujo a un aumento tanto de la producción de C-1027 como heptaeno. Además, la sobreexpresión de *sgcR1* en la cepa mutante Δ *sgcR* condujo a un aumento de siete veces aproximadamente de la producción de C-1027. *sgcR3* ocupa una regulación de nivel superior mediante el

control de sgcR1 y sgcR2 en la regulación de jerarquía de la producción de C-1027. En conclusión, la alteración y los efectos de expresión de orf24 y la comparación de Orf24 con otros ortólogos caracterizados funcionalmente indican que Orf24 actúa como un regulador/activador positivo específico de la vía en la producción de enduracidina.

Orf18 es un putativo regulador de respuesta huérfano atípico y se alinea con ortólogos funcionalmente confirmados

La producción de antibióticos en las especies de *Streptomyces* está estrechamente regulada por redes genéticas complejas que limitan la capacidad de muchos productores de antibióticos de tipo silvestre para generar los rendimientos necesarios para una producción industrial a gran escala y rentable. Un mecanismo regulador importante son los sistemas de transducción de señales de dos componentes. Los sistemas de dos componentes incluyen un sensor de quinasa y un regulador de respuesta afín. El sensor de quinasa responde a estímulos/señales ambientales externas específicas, tales como estrés, nutrición y sustancias químicas, etc., y luego transmite la señal a un regulador de respuesta citoplásmica que desencadena y activa la transcripción de los genes diana. Un regulador de respuesta que no está emparejado con un sensor de quinasa se designa como un regulador de respuesta huérfano.

Los sistemas de dos componentes y los reguladores de respuesta huérfanos están presentes en los genomas de estreptomicetos y pueden funcionar para reprimir la producción de metabolitos secundarios. En el grupo de genes de enduracidina de *S. fungicidicus*, *orf18* codifica un supuesto regulador de respuesta huérfana putativa que comparte una similitud de secuencia baja a moderada con otros tres reguladores de respuesta de *Streptomyces* caracterizados, incluido un regulador de respuesta huérfana, SCO3818, de *S. coelicolor* (FIG. 12). Orf18 tiene una secuencia de extremo N-terminal más larga en comparación con las otras proteínas alineadas y parece ser un regulador de respuesta huérfano atípico porque una lisina altamente conservada en la posición 118 (relativa a la posición común 105) está ausente en Orf18 y se reemplaza con una treonina. Se propone que la lisina sea necesaria para formar el bolsillo de fosforilación.

Sólo unos pocos reguladores de respuesta a estreptomicetos han sido funcionalmente caracterizados. El genoma de *S. coelicolor* contiene un total de cinco reguladores de respuesta huérfanos atípicos y siete típicos. Orf18 comparte 26% de identidad aa en 191 aa superpuestas con AbsA2. La eliminación de AbsA2 en *S. coelicolor* resultó en un aumento

de la producción de dos antibióticos, actinorhodina y undecil-prodigiosina. Orf18 muestra 32% de identidad aa en 176 aa superpuestas con SCO3818. La eliminación de sco3818 condujo a una producción mejorada de actinorodina. Orf18 comparte un 29% de identidad aa en 166 superpuestas aa con SCO1745 (AbrA2). La eliminación del operón regulador de respuesta que contiene AbrA2 dio como resultado un aumento del 100% del antibiótico antitumoral de oviedomicina en la cepa recombinante *S. coelicolor* M145 en comparación con el productor de tipo silvestre. El papel regulador negativo observado de Orf18 en la producción de enduracidina es consistente con las actividades demostradas de los reguladores negativos relacionados (FIG. 12). Además, se observa que Orf18 comparte la mayor similitud de secuencia de proteínas con los miembros de la familia LuxR de reguladores transcripcionales en la búsqueda de BLAST.

Ausencia de efectos polares en el mutante BM38-1.orf18pfrd-AmR

La región eliminada en la cepa BM38-1.18pfrd-AmR mutante involucra tres genes, orf18, la región que codifica la porción de extremo N-terminal de orf17 localizado corriente abajo de orf18, y la totalidad del orf19 ubicado corriente arriba de orf18 (FIGS. 5 y 10). Se predice que orf17 codifica una ribonucleasa que aparentemente no tiene ninguna función relacionada con la biosíntesis o regulación de la enduracidina. Además, el gen de resistencia a la apramicina que reemplaza a *orf18* y su región flanqueante se transcribe de forma divergente con *orf17* y no debe crear ningún evento de lectura a través del promotor del gen de resistencia a apramicina. Por lo tanto, no debería haber efectos polares resultantes de la eliminación parcial de *orf17*.

orf19 se transcribe y traduce en la misma dirección que *orf18*. Este gen está anotado para codificar una proteína de función desconocida. La cepa mutante SfpXYF24D3 que porta la alteración de orf18 sola y el mutante BM38-1.18pfrd-AmR que porta la eliminación de *orf18* y *orf19* juntos tienen efectos similares en la producción de enduracidina, lo que implica que *orf19* no tiene ningún papel o un papel insignificante en la producción de enduracidina. El gen orf20 se ubica corriente arriba de *orf19* y se transcribe y traduce en la misma dirección que el marcador de resistencia a apramicina insertado (FIG. 10) orf20 sigue intacto en BM38-1.18pfrd-AmR y el producto aparentemente no tiene un papel en la producción de enduracidina. Por lo tanto, no se cree que ningún efecto polar sobre la expresión de orf20 sea responsable de la producción mejorada de enduracidina en BM38-1.18pfrd-AmR.

Ejemplo 7

Aplicaciones y manipulaciones adicionales de *orf24* y/o *orf18* para mejorar las cepas productoras de enduracidina

Además de los ejemplos proporcionados anteriormente, hay otras formas posibles de utilizar los roles reguladores de *orf24* y *orf18* para mejorar la producción de enduracidina.

i. Expresión de *orf24* bajo un promotor de sobreexpresión alternativo, constitutivo o inducible

pXY152-endorf24 (mostrado en la FIG. 2) se construyó para la expresión ectópica integradora de *orf24* bajo el control de ermE*p, un estreptomiceto ampliamente utilizado como el promotor de la expresión constitutivo fuerte. La sobreexpresión de *orf24* también puede ser impulsada por otros promotores constitutivos o inducibles. El promotor tipA es un promotor de estreptomicetos de sobreexpresión inducible por tiostrepton. Se desarrolló un plásmido de lanzadera de *E. coli-Streptomyces* con el promotor tipA multicopia, pXY200, que se ha utilizado con éxito para la sobreexpresión de genes de estreptomicetos. Para aplicaciones relevantes a esta descripción, el promotor tipA puede escindirarse de pXY200 y clonarse en pXY152 para reemplazar ermE*p y dirigir la expresión de *orf24*. Del mismo modo, *orf24* puede transferirse fácilmente de pXY152-endorf24 a pXY200 para la expresión basada en plásmidos. Otras opciones de promotores incluyen, pero sin limitarse a, el sistema P(nitA)-NitR y el promotor de estreptomicetos SF14. Recientemente, el plásmido integrador pKC1139 y el promotor nativo del gen resistente a la apramicina se utilizaron con éxito para expresar genes reguladores para la hiperproducción del antibiótico peptídico teicoplanina. El gen regulador sanG codifica un activador específico a la vía para la producción de nikomicina. La expresión de una copia adicional de sanG bajo el control de cinco promotores diferentes (P_{hrdB} , P_{tcp830} , P_{SF14} , P_{ermE^*} y P_{neos}) llevó a aumentos en los rendimientos de nikomicina en un 69%, 51%, 26%, 22% y 13%, respectivamente (véase Du, Applied Microbiology and Biotechnology 97: 6383-6396, 2013).

ii. Cepas mutantes dobles de *S. fungicidicus* con eliminación de *orf18* y sobreexpresión de *orf24*

Tanto con el mutante de delección *orf18* como con las cepas de sobreexpresión *orf24* que exhiben una mayor producción de enduracidina, se puede generar un doble mutante que contenga ambos y si se observa un efecto aditivo sobre el rendimiento de este antibiótico peptídico. Se puede crear el mutante doble introduciendo el plásmido de sobreexpresión

pXY152-endorf24-blatsr (FIG. 8) en el mutante BM38-2.18pfrd-AmR. pXY152-endorf24-blatsr es un plásmido de integración conyugal que porta un gen de resistencia a la tiosstrepton (tsr) para la selección en estreptomicetos y el gen de resistencia a la ampicilina (bla) para la selección en *E. coli*. Debido a que la cepa de *E. coli* S17-1 utilizada para la conjugación es naturalmente resistente al cloranfenicol (leva), el marcador de resistencia al cloranfenicol en pXY152-endorf24-camtsr (véase más arriba) ha sido reemplazado por resistencia a la ampicilina (bla) para seleccionar S17-1 transformantes. Alternativamente, pXY152-endorf24-camtsr y sus derivados pueden introducirse en estreptomicetos utilizando una cepa de *E. coli* conyugal diferente, ET12567/pUZ8002.

Usando bien sea el plásmido pXY152-endorf24-blatsr o pXY152-endorf24-camtsr para introducir la segunda copia de orf24 en el mutante orf18 deficiente, es posible seleccionar el mutante doble mediante la resistencia a tiosstreptona. Para generar un mutante orf18 nulo con eliminación en el marco en BM38-2 (ATCC PTA-122342), los plásmidos pXY300-orf18ifd (FIG. 3) y pKS-orf18ifd-T-AmR (NS) (FIG. 6) fueron contruidos con este fin. pXY300-orf18ifd permite la selección del mutante de deleción en el marco *orf18* con tiosstreptona mientras PKS-orf18ifd-T-AmR (NS) utiliza apramicina para seleccionar mutantes de deleción en el marco. Aunque las cepas mutantes de *S. fungicidicus* de tipo silvestre se seleccionan fácilmente usando el marcador de resistencia tiosstreptona, se han encontrado dificultades al usar este marcador de resistencia en la cepa BM38-2 (ATCC PTA-122342). Por lo tanto, dos plásmidos, pXY300-orf18ifd y pKS-orf18ifd-T-AmR (NS), se construyeron para el mismo propósito.

En vista de todas las modalidades posibles a las que se pueden aplicar los principios de la invención descrita, debe reconocerse que las modalidades ilustradas son solo ejemplos preferidos de la invención y no deben tomarse como limitantes del alcance de la invención. Más bien, el alcance de la invención se define mediante las siguientes reivindicaciones. Por lo tanto, reivindicamos como nuestra invención todo lo que está dentro del alcance y el espíritu de estas reivindicaciones.

Reivindicaciones:

1. Una cepa recombinante de *Streptomyces fungicidicus* que comprende uno o más genes modificados seleccionados del grupo que consiste en un marco de lectura abierta aumentada-24 (*orf24*), un marco de lectura abierta disminuida-18 (*orf18*) y tanto una *orf24* aumentada como una *orf18* disminuida; en donde se obtiene una producción mejorada de enduracidina con la cepa recombinante de *Streptomyces fungicidicus* en comparación con la obtenida con una cepa de control de *Streptomyces fungicidicus*.
2. La cepa recombinante de la reivindicación 1, en donde el *orf18* disminuido disminuye porque se ha anulado.
3. La cepa recombinante de la reivindicación 2, en donde el *orf18* disminuido se ha anulado mediante un proceso seleccionado del grupo que consiste en una eliminación en el marco, una mutación de cambio de marco, una mutación puntual y cualquier combinación de las mismas.
4. La cepa recombinante de la reivindicación 3, en donde el *orf18* disminuida se ha anulado mediante una eliminación en el marco.
5. La cepa recombinante de la reivindicación 4, en donde la eliminación en el marco es de los nucleótidos 5 a 660 de *orf18* (SEC ID NO: 27).
6. La cepa recombinante de cualquiera de las reivindicaciones 1-5, en donde el *orf24* aumentado está unido operativamente a un promotor heterólogo.
7. La cepa recombinante de la reivindicación 6, en donde el promotor heterólogo es un promotor constitutivo fuerte.
8. La cepa recombinante de la reivindicación 7, en donde el promotor constitutivo fuerte es *ermE**p.

9. La cepa recombinante de cualquiera de las reivindicaciones 1-8, en donde el ORF24 aumentado se aumenta porque se ha sobreexpresado.

10. La cepa recombinante de cualquiera de las reivindicaciones 1-9, en donde el *Streptomyces fungicidicus* es *Streptomyces fungicidicus* ATCC 21013.

11. La cepa recombinante de cualquiera de las reivindicaciones 1-9, en donde el *Streptomyces fungicidicus* es *Streptomyces fungicidicus* ATCC PTA-122342.

12. La cepa recombinante de cualquiera de las reivindicaciones 1-11; en donde la producción de enduracidina mediante la cepa recombinante es al menos 1.2 veces mayor que la producción de enduracidina por la *Streptomyces fungicidicus* de control.

13. La cepa recombinante de cualquiera de las reivindicaciones 12; en donde la producción de enduracidina mediante la cepa recombinante es al menos de 1.2 a 4.6 veces mayor que la producción de enduracidina por la *Streptomyces fungicidicus* de control.

14. La cepa recombinante de la reivindicación 1, que es BM38-2.24/16 (Nº de depósito ATCC PTA-124006).

15. La cepa recombinante de la reivindicación 1, que es BM38-2.18pfrd-AmR (Nº de depósito ATCC PTA-124007).

16. Un método para producir enduracidina, que comprende cultivar la cepa recombinante de *Streptomyces fungicidicus* de cualquiera de las reivindicaciones 1-15 bajo condiciones suficientes para producir enduracidina.

17. El método de la reivindicación 16, que comprende además aislar la enduracidina del medio de cultivo.

18. Un vector de interrupción que comprende un marco de lectura abierta *orf18* (SEC ID NO: 37) el cual ha sido anulado.

19. El vector de interrupción de la reivindicación 18, en donde el *orf18* ha sido *anulado* mediante un proceso seleccionado del grupo que consiste en una eliminación en el marco, una mutación de cambio de marco, una mutación puntual y cualquier combinación de las mismas.

20. El vector de interrupción de la reivindicación 19, en donde el *orf18* ha sido anulado por una eliminación en el marco.

21. El vector de interrupción de la reivindicación 20, en donde la eliminación en el marco es de los nucleótidos 5 a 660 de *orf18* (SEC ID NO: 27).

22. Un vector de expresión que comprende un marco de lectura abierta aumentada *orf24* (SEC ID NO: 38).

23. El vector de expresión de la reivindicación 22, en donde el *orf24* aumentado está unido operativamente a un promotor heterólogo.

El vector de expresión de la reivindicación 23, en donde el promotor heterólogo es un promotor constitutivo fuerte.

25. El vector de expresión de la reivindicación 24, en donde el promotor constitutivo fuerte es *ermE**p.

26. El vector de expresión de la reivindicación 22, en donde el vector de expresión se selecciona del grupo que consiste de pXY152-endorf24 (SEC ID NO: 3), pXY152-endorf24-camtsr (SEC ID NO: 20), y pXY152-endorf24-blatsr (SEQ ID NO: 23).

27. El vector de interrupción de la reivindicación 18, seleccionado del grupo que consiste en pXY300-orf18ifd (SEC ID NO: 8), pKS-T-orf18ifd (SEQ ID NO: 11), pKS-T-orf18pfrd-AmR (SEQ ID NO: 14), y pKS-orf18ifd-T-AmR(NS)(SEQ ID NO: 19).

COMPOSICIONES Y MÉTODOS PARA LA PRIDUCCIÓN INCREMENTADA DE ENDURACIDINA EN UNA CEPA GENÉTICAMENTE MODIFICADA DE STREPTOMYCES FUNGICIDICUS

RESUMEN DE LA DIVULGACIÓN

Esta descripción describe composiciones y métodos para mejorar la producción de enduracidina en cepas genéticamente modificadas de *Streptomyces fungicidicus*. En particular, la presente descripción describe la manipulación genética de los genes reguladores *orf24* y *orf18* asociados con la agrupación de genes de biosíntesis de enduracidina (enramicina) de *Streptomyces fungicidicus* para generar constructos de vectores y cepas recombinantes que producen mayores rendimientos de enduracidina.

LISTADO DE SECUENCIAS

<110> Oregon State University
Zabriskie, T. Mark
Yin, Xihou

<120> COMPOSICIONES Y MÉTODOS PARA LA PRODUCCIÓN INCREMENTADA DE ENDURACIDINA
EN UNA CEPA GENÉTICAMENTE MODIFICADA DE STREPTOMYCES FUNGICIDICUS

<130> 127789-236560-P005PCT

<150> 62/430,838

<151> 2016-12-06

<150> 62/479,087

<151> 2017-03-30

<160> 40

<170> PatentIn version 3.5

<210> 1

<211> 31

<212> ADN

<213> Secuencia Artificial

<220>

<223> Cebador de Oligonucleótido Sintético

<400> 1

ccaccacata tggaaataag ttctgtctcc a

31

<210> 2

<211> 32

<212> ADN

<213> Secuencia Artificial

<220>

<223> Cebador de Oligonucleótido Sintético

<400> 2

gtgtgtgaat ttctgttca ccggccaga tg

32

<210> 3

<211> 6210

<212> ADN

<213> Secuencia Artificial

<220>

<223> Pl 痘mido Sint 驛 ico pXY152-endorf24

<400> 3

gaattcgtaa tcatgtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca	60
cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaata agtgagctaa	120
ctcacattaa ttgcgttgcg ctcaactgccc gctttccagt cgggaaacct gtctgtccag	180
ctgcattaat gaatcggcca acgogcgggg agaggcggtt tgcgtattgg gcgctcttcc	240
gcttctctgc tcaactgactc gctgcgctcg gtcttccggc tgcggcgagc ggtatcagct	300
cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg	360
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc	420
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga	480
aaccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct	540
cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg	600
gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag	660
ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttata cggttaactat	720
cgtcttgagt ccaaccoggt aagacacgac ttatcgccac tggcagcagc cactggtaac	780
aggattagca gagcgaggta ttaggcgggt gctacagagt tcttgaagtg gtggcctaac	840
tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttaccttc	900
ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggttt	960
tttgtttgca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga tcctttgatc	1020
ttttctacgg ggtctgacgc tcagtgaac gaaaactcac gttaagggat tttggtcatg	1080
agattatcaa aaaggatctt cacctagatc cttttggttc atgtgcagct ccatcagcaa	1140
aaggggatga taagtttata accaccgact atttgcaaca gtgccgttga tcgtgctatg	1200

atcgactgat gtcacacg gtaggtgca atgtcgtgca atacgaatgg cgaaaagccg	1260
agctcatcgg tcagcttctc aaccttgggg ttacccccgg cgggtgtgtg ctggtccaca	1320
gctccttccg tagcgtccgg cccctcgaag atgggccact tggactgac gaggcctgc	1380
gtgctgcgct ggggtccggga gggacgctcg tcatgccctc gtggtcaggt ctggacgacg	1440
agccgttcga tcttgccacg tcgcccgtta caccggacct tggagttgtc tctgacacat	1500
tctggcgcct gccaaatgta aagcgcagcg cccatccatt tgcctttgcg gcagcggggc	1560
cacaggcaga gcagatcac tctgatccat tgcccctgcc acctcactcg cctgcaagcc	1620
cggctgccccg tgtccatgaa ctcgatgggc aggtacttct cctcggcgtg ggacacgatg	1680
ccaacacgac gctgcatctt gccgagttga tggcaaaggt tccctatggg gtgccgagac	1740
actgcacat tcttcaggat ggcaagttgg tacgcgtcga ttatctcgag aatgaccact	1800
gctgtgagcg ctttgccttg gcggacaggt ggctcaagga gaagagcctt cagaaggaag	1860
gtccagtcgg tcatgccttt gctcggttga tccgctcccg cgacattgtg gcgacagccc	1920
tgggtcaact gggccgagat ccgttgatct tcttgcatcc gccagaggcg ggatgcgaag	1980
aatgcgatgc cgctcgccag tcgattggct gagctcatga gcggagaacg agatgacgtt	2040
ggaggggcaa ggtcgcgctg attgctgggg caacacgtgg agcggatcgg ggattgtctt	2100
tcttcagctc gctgatgata tgctgacgct caatgccgtt tggcctccga ctaacgaaaa	2160
tccgcattt ggacggctga tccgattggc acggcggacg gcgaatggcg gagcagacgc	2220
tcgtccgggg gcaatgagat atgaaaaagc ctgaactcac cgcgacgtat cgggccctgg	2280
ccagctagct agagtgcacc tgcaggtccc cggggatcgg tcttgacctg ctgctcgtg	2340
atgtacttca ccagctccgc gaagtcgctc ttcttgatgg agcgcattgg gacgtgcttg	2400
gcaatcacgc gcacccccg gccgttttag cggctaaaaa agtcatggct ctgccctcgg	2460
gcggaccacg cccatcatga ccttgccaag ctcgctctgc ttctcttcga tcttcgccag	2520
cagggcgagg atcgtggcat caccgaaccg cgccgtgcgc gggtcgtcgg tgagccagag	2580
tttcagcagg ccgccaggc ggcccaggtc gccattgatg cgggccagct cgcggacgtg	2640

ctcatagtcc acgacgccc tgattttgta gccctggccg acggccagca ggtaggccga	2700
caggctcatg ccggccgcc cgcccttttc ctcaatcgct cttcgttcgt ctggaaggca	2760
gtacaccttg ataggtgggc tgcccttcct ggttggcttg gtttcatcag ccacccgctt	2820
gccctcatct gttacgccg cggtagccgg ccagcctcgc agagcaggat tcccgttgag	2880
caccgccagg tgcaataag ggacagtga gaaggaacac ccgctcgcgg gtgggcctac	2940
ttcacctatc ctgccggct gacgccgttg gatacaccaa ggaaagtcta cacgaaccct	3000
ttggcaaaat cctgtatatc gtgcgaaaaa ggatggatat accgaaaaa tcgctataat	3060
gacccgaag cagggttatg cagcggaaaa gatccgtcga cctgcaggca tgcaagctct	3120
agcgattcca gacgtccga aggcgtggcg cggcttcccc gtgccggagc aatcgccctg	3180
ggtgggttac acgacgcccc tctatggccc gtactgacgg acacaccga gccccggcgg	3240
caaccctcag cggatgcccc ggggcttcac gttttccag gtcagaagcg gttttcggga	3300
gtagtgcccc aactggggta acctttgagt tctctcagtt gggggcgtag ggtcgccgac	3360
atgacacaag gggttgtgac cgggggtggac acgtacgcgg gtgcttacga ccgtcagtcg	3420
cgcgagcgcg agaactcgag cgcagcaagc ccagcgacac agcgtagcgc caacgaagac	3480
aaggcggccg accttcagcg cgaagtcgag cgcgacgggg gccggttcag gttcgtcggg	3540
catttcagcg aagcgccggg cacgtcggcg ttccggacgg cggagcgccc ggagttcgaa	3600
cgcacctga acgaatgccg cgcggggcgg ctcaacatga tcattgtcta tgacgtgtcg	3660
cgtttctcgc gcctgaaggt catggacgcg attccgattg tctcggaatt gctcgccctg	3720
ggcgtgacga ttgtttccac tcaggaaggc gtcttcgggc agggaaacgt catggacctg	3780
attcacctga ttatgcggct cgacgcgtcg cacaaagaat cttcgtgaa gtoggcgaag	3840
attctcgaca cgaagaacct tcagcgcgaa ttgggcgggt acgtcggcgg gaaggcgcct	3900
tacggcttcg agcttgtttc ggagacgaag gagatcacgc gcaacggccg aatggtcaat	3960
gtcgtcatca acaagcttgc gactcgacc actcccccta ccggaccctt cgagttcgag	4020

cccgcacgtaa tccggtggtg gtggcgtgag atcaagacgc acaaacacct tcccttcaag	4080
cggggcagtc aagccgccat tcacccgggc agcatcacgg ggctttgtaa gcgcatggac	4140
gctgacgccg tgccgacccg gggcgagacg attgggaaga agaccgcttc aagcgcctgg	4200
gacccggcaa ccgttatgcg aatccttcgg gacccgcgta ttgcgggctt cgccgctgag	4260
gtgatctaca agaagaagcc ggacggcacg ccgaccacga agattgaggg ttaccgcatt	4320
cagcgcgacc cgatcacgct ccggccggtc gagcttgatt gcggaccgat catcgagccc	4380
gctgagtggc atgagcttca ggcgtggttg gacggcaggg ggcgcgcaa ggggctttcc	4440
cgggggcaag ccattctgtc cgccatggac aagctgtact gcgagtgtgg cgccgtcatg	4500
acttcgaagc gcggggaaga atcgatcaag gactcttacc gctgccgtcg ccggaaggtg	4560
gtcgaccgt ccgcacctgg gcagcacgaa ggcacgtgca acgtcagcat ggcggcactc	4620
gacaagttcg ttgcggaacg catcttcaac aagatcaggc acgccgaagg cgacgaagag	4680
acgttggcgc ttctgtggga agccgcccga cgcttcggca agctcactga ggcgccctgag	4740
aagagcggcg aacgggcgaa cttgtttgcg gagcgcgcgg acgccctgaa cgoccttgaa	4800
gagctgtacg aagaccgcgc ggcaggcgcg tacgacggac ccgttggcag gaagcacttc	4860
cggaagcaac aggcagcgct gacgctccgg cagcaagggg cggaagagcg gcttgccgaa	4920
cttgaagccg ccgaagcccc gaagcttccc cttgaccaat ggttccccga agacgccgac	4980
gctgaccga ccggccctaa gtcgtggtgg gggcgcgcgt cagtagacga caagcgcgtg	5040
ttcgtcgggc tcttcgtaga caagatcgtt gtcacgaagt cgactacggg cagggggcag	5100
ggaacgcca tcgagaagcg cgcttcgac acgtgggcga agccgccgac cgacgacgac	5160
gaagacgacg ccgaggacg cacggaagac gtagcggcgt aggtgagtga agatctagac	5220
gtggcaccgc gatgctgttg tgggcacaat cgtgccggtt ggtaggatcc acatatggaa	5280
ataagttcgc tctccaccga cggctccccg cggatcgacg gggagagtcc cgagcacgtg	5340
gaaatgctgg ccgccgccga caccgcgctt ccaccgatca tggcgcaccg ccgcaccggg	5400
cgggtcatcg acggcatgca ccggctgcgc gccgcgatgc tgacgggccc tacgacgac	5460

gcggtgaggt tcttcgacgg caccgaggag gacgccttcg tcctcgccgt gaagtcgaac	5520
atcgcgcacg gactgccgct gtccgcccgc gaccgccggc gggccgccgg gcgcatcatg	5580
gccacccatc cccggtggtc ggaccggatg atgcctcgg tggtcggcac ctccgccagg	5640
acggtcgccg agatccgccg cgacgccggc gccgccgggg cgggggagcc caccgcac	5700
ggccgggacg gcagggtacg gcccgtcgac gtgagcgagg gccgcagact ggcccacgac	5760
atgatcgtcc gcgaccggg cctgtcgtc cgccaggctc cccgcgccgc cgggatctcg	5820
ccggagaccg tcagggacgt cagacaccgg atgctccgc gtgaggacc ggtgccgcg	5880
ccgcggccgc ggaccctggt ggagcgccgc gcggaccgcc gggcggagcc ggccgggaag	5940
gccgccgcgc cgtgcgggac ggagccgcc cccgccgtc tgatgaagcg gctgagggcc	6000
gatccggcgc tgcgtctcaa cgagaacgga cgcgacctgc tgcggcttct ggatatccac	6060
acggtccggc tggaggactg gaaccgcatt atcgaagcg tgccgccga ccgtctggag	6120
acggtggcgc agctggcacg ctctcgccc gacaaatggt ccgagatgc gtcacgcac	6180
gaaagcaacg catcacatct ggccgggtga	6210

<210> 4
 <211> 38
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 4 ttattgaagc ttgccggggc cgacgcggcg ggccgcct	38
--	----

<210> 5
 <211> 34
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 5
gttgttttaa ttaaacacca ggcctcctgg ggtg 34

<210> 6
<211> 38
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 6
tttatattaa ttaatgaccc ttccgtcccg cccccgat 38

<210> 7
<211> 30
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 7
tttggtgcta gctggtcgtg gcgctgttcc 30

<210> 8
<211> 10670
<212> ADN
<213> Secuencia Artificial

<220>
<223> Pl 痘 mido Sint 騁 ico pXY300-orf18ifd

<400> 8
ccgcggctcc tgcccgccga acgcgtcgtc gtcgacggcc tgggtgctcat cgacgagcac 60

ccggagccag gtgaaaagcg ccggcggacg ctcgactgg gcgcgggatt ccagcagtaa 120

cccaggtccg ccggccacct caccgcaggc agaccctcgg ctttcgcgcc gggaccgcca 180

tgagaccgcc acccggtatgt ccgggggtggc ggtctcatgg cgtccctca gcggccctac 240

gcgaccgctg tgtcgacgcg gaggcagtct ccggggagcc ggccccaggg ctggagtacc 300

ggggcgagct tgcccttgca gcggcggcac acgggcgacg cggcgccggg cgggggagtc	360
ttgacgtcga cggtgaccgg cttcggcttc aggcgcttcc gaaggctgat cgtcgggcgg	420
atgcctcct tcttcggctg acgcactcgg tcgagcgaag cggccagctc ctccgtgcgg	480
gcctcgttgg ccttcggcg cgcctcggg aaagcagctt cctcctcgga catgacctcg	540
aacctcatct ggtcagcgtc aaggctgccc ggcgggccg gcgcttcgg gcggggcggg	600
tccaggacgt ccttgcccca caccaggccc caggactcga cgagccgccg gacgcccggt	660
aggccgtacg tctcggcgac cttgatgagg tcgaggtcga ggcgacgtcc ggcgacgcgg	720
gcgatgtatc ggtaccagat gtaggcggg atgaccgga tggcgaccag gccctcggtg	780
tcgtcgggta tctcctctc ggtgcggacg tcctgctgga tgccgagttc cttgatcagc	840
cggttcaggt tctgcgaccg gtagtgcttg cggacctgga agacgccgaa ctgcgctcg	900
cgggtacttct cgacgaacgg gccggggcca cgaagccgt gcagctcggc ggccgccgcg	960
tcgccaggt cgagcggtc catgcggtcg tcgccgcgac cggccttgaa gttctgtccg	1020
gccagctcca ggccgatctt ggcgacgcc cccttggtct tgtcgccgtc cttgtagagg	1080
tagcgggcct gcttgccgc atcgccgtca gcggcgccg gccggttag tggcgcacg	1140
tcggtgccgt ggcccttgcc ctacacaggag caaccgggc ggtcgacgt ctgcgtgacg	1200
gtgtagccgc ccgcggattc gacccggcg gccaggctc cggcgagtgc gtgcgggaa	1260
gcctgggcgt ccgggccgag cacctcggg gtgaccaga gcgtgtgcca gtgcaggtc	1320
cagccggagc ccagccgaa ggtgtcctcg aaggccgct cgtagccgat gatccgaag	1380
tcgtcgcga tcgtgcgac gcggcggccg gacgagcgt acgcgccctt ccagccgtcg	1440
tgcaagaccg cgaccaggcc gtgcgcatt cccttgcgga cggtgccgaa cgocatgcg	1500
tcgaagtggc gcaacgtgtt cgtgccaagg tgcagccgt acccggcgtc gcgagaccg	1560
tcggcggcga gctgcacgtt cgagccccgt acggccagga tgcggctcat gcaccacggg	1620
cagggtgtgga cgttgttgca gcggcacgtg ttgccccacg tcgcctcgcc cggcttcac	1680

atcagctcgc ggtccggcag tgagccgggt cccgcagccc ttgaacgcct cgttcagcga	1740
caccgtctgg tgccggtccc gccgggcgaa ccgctcgtcg cggcgggtcc tcccgcccgc	1800
tgtcgcggca ccctcgtttg gggtagaacc cgttcagtt acagcgctct gacctgcagt	1860
ggacggagat tttcccttac tactaaagcc cgcgtccgga ttaccgcgtg tagtcgtgct	1920
tgctacgctg cgtgactggt ccgcaatgag acgctttgcg cgctttcggc aggcgctcga	1980
gcagtagatt ttggggcgct tcccggggat gtggacgac ggggtgccgc agtggcactt	2040
cggctccggcg gggcgcggtg gtgtcgacgc gctgttctct cgtacgcctc gtcacagagc	2100
aaacgtcctc actcggcatg ctgcgcgggt tcgggggcgg cgagccggga ggccaatccc	2160
gggctcgtgc catttctggg tcctgttgat catcactgac gaatcgaggt cgaggaaccg	2220
agcgtccgag gaacagaggc gcttatcggt tggccgcgag attcctgtcg atcctctcgt	2280
gcagcgcgat tccgagggaa acggaaacgt tgagagactc ggtctggctc atcatgggga	2340
tggaaccga ggcggaagac gcctcctcga acaggtcggga agggccacc ttttcgctgc	2400
cgaacagcaa ggccagccga tccggattgt ccccgagttc cttcacggaa atgtcgccat	2460
ccgccttgag cgtcatcagc tgcataccgc tgtcccgaat gaaggcgatg gcctcctcgc	2520
gaccggagag aacgacggga agggagaaga cgtaacctcg gctggccctt tggagacgcc	2580
ggtccgcgat gctggtgatg tcaactgtcga ccaggatgat ccccgacgct ccgagcgcga	2640
gcgacgtcgc tactatcgcg ccgatgttcc cgacgatctt caccctcgtc agaacgacga	2700
cgtccccacg ccggctcgcg atatcgccga acctggccgg gcgagggacg cgggcgatgc	2760
cgaatgtctt ggcccttcgc tcccccttga acaactgggt gacgatcgag gattcgatga	2820
ggcggaccgg tatgttctgc cggccgcaca gatccagcaa ctcatatgga aaaggactgc	2880
tgtcgtgcc gtagacctcg atgaactcca ccccgccgc gatgctgtgc atgaggggct	2940
cgacgtcctc gatcaacgtt gtctttatgt tggatcgga cggcttggtg acatcgatga	3000
tccgctgcac cgcgggatcg gacggatttg cgatggtgtc caactcagtc atggtcgtcc	3060
taccggctgc tgtgttcagt gacgcgattc ctgggggtgtg acaccctacg cgacgatggc	3120

ggatggctgc cctgaccggc aatcaccaac gcaaggggaa gtcgtcgctc tctggcaaag	3180
ctccccgctc ttccccgtcc gggaccgcg cggtcgatcc ccgcatatgg tgcactctca	3240
gtacaatctg ctctgatgcc gcatagttaa gccagccccg acaccggcca acaccgctg	3300
acgcgccctg acgggcttgt ctgctcccg catccgtta cagacaagct gtgaccgtct	3360
cggggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg agacgaaagg	3420
gcctcgtgat acgcctatit ttataggtta atgtcatgat aataatggtt tcttagacgt	3480
caggtggcac ttttcgggga aatgtgcgcg gaaccctat ttgtttatit ttctaaatac	3540
attcaaatac gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa	3600
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattccctit ttgcggcac	3660
tttgccctcc tgtttttgct caccagaaa cgctggtgaa agtaaaagat gctgaagac	3720
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga	3780
gttttcgccc cgaagaacgt ttccaatga tgagcactit taaagtictg ctatgtggcg	3840
cggatattac ccgtattgac gccgggcaag agcaactcgg tcgccgata cactattctc	3900
agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag	3960
taagagaatt atgcagtgtc gccataacca tgagtataa cactgcccgc aacttacttc	4020
tgacaacgat cggaggaccg aaggagctaa ccgctttit gcacaacatg ggggatcatg	4080
taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg	4140
acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact ggogaactac	4200
ttactctagc ttcccgcaa caattaatag actggatgga ggcggataaa gttgcaggac	4260
cacttctgcg ctccggccct ccggctggct ggtttattgc tgataaatct ggagccggtg	4320
agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg	4380
tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg	4440
agataggtgc ctactgatt aagcatttgt aactgtcaga ccaagtttac tcatatatac	4500

tttagattga tttaaaactt catttttaat ttaaaaggat ctaggtgaag atcctttttg	4560
ataatctcat gacaaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg	4620
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc	4680
aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc	4740
tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt	4800
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctgctctgc	4860
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact	4920
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt tcgtgcacac	4980
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag	5040
aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaacg ggcagggtcg	5100
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg	5160
togggtttcg ccacctctga cttgagcgtc gatTTTTgtg atgctcgtca ggggggcgga	5220
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt	5280
ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct	5340
ttgagtgagc tgataccgct gcgcgcagcc gaacgaccga gcgcagcgag tcagtgagcg	5400
aggaagcgga agagcgccca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt	5460
aatgcagctg gcacgacagg ttcccgact ggaaagcggg cagtgagcgc aacgcaatta	5520
atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta	5580
tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt	5640
acgccaagct tgccggggcc gacgcggcgg gcggcctcgt ccatcgggac gcgccagtcg	5700
acgccgacga tgtccgcgcc ggctcgcgcc atgagcccca gcagctcgcc ggtgccgacg	5760
ccgaagtgga tgccggggac gccgtggccg gccaccgcgc ggaacacctt cgcgcaggcg	5820
ggcagcaccg aacgccggta gtcggagggg gcgagcgcgc cggcccagga gtogaagagc	5880
tgcacggccg aggcgccggc ccggatctgg acgtcgagga aggcgcgccgt gatgtcggcg	5940

aggcggtcga gcaggtcggc ccagagctcg gggtcgccgt acatcatcgc cttggcgttc	6000
tcgtacgtac gggacggggc gccctcgacg aggttaactcg caagggtgaa cggggcgccc	6060
gcgaaaccga tcagcggggt ggacccgagc tcacgggtca gcatgccgat ggcctcggtg	6120
acgtaggaga cgtcctccgg ggtcaggtcg cgcagccggg cgaggtcggc gcgggtgcgc	6180
accggctgct cgacgaccgg gccgatgccg ggcttgatgt cgaggtcgat gccgatggcc	6240
ttgagcggga cgacgatgtc gctgaagtag atcgccgcgt cgacgtggtg gcggcgcacc	6300
ggctggaggg tgatctcggt gaccagctcg ggccgcatgc aggagtcgag catcccgatg	6360
ccctcgcgca ccttcgggta ctccggcagt gagcgcccg cctgccgat gaaccacacc	6420
ggcgtgtgcg gcaccggctc gcgccggcac gccttgagga aggctgattc ccggacggcg	6480
tcgttcttga cggggtcggg ggtcgcggta ggcgtctggc ctttcgggct cgtgttgga	6540
ctcacaccgg ccagtctcgc acgaccgctc cgccgccccg gacccaggg gcgtccggcg	6600
ggcggggccg cgggtcagga tccggacagc gccgggcggc gcgggtgtcc ctccctgcgc	6660
cgggggcccc ttccgcttaa tcttccggc atggctgcgg ctcagggacg actgtcggac	6720
ggcgctggcg gaatggacga accgaaggag ggcggggggg atcccgggca cggaggtgcg	6780
cctccgccgc ctttcgggc cgctgtcgag gcgtgcaga gcgccggct gcggccgcag	6840
atcgaggtgg agacggtgcc cgcgccgaaa cggctcgccc cgtacgcgca cgcgctggag	6900
gcggcggtcg tcgacggcga ggaggatctg gccgacggc ggctggtgct gctgtgcgac	6960
ccggccggac acgacgcctg gcgggggacc ttccgtctgg tgacgtggt gcgcgccgag	7020
ctggagccgg agatggcggc ggatccgctg ctgccggacg tgtgctggtc ctggctgacc	7080
ggcgcgctgg cggcgcgcg cctgtcgtac ggcgagccga gcggcacggt gacgcgggcg	7140
agttcgcaact acttcggcgg gctgtccgcg cggcccgcg cctcccagat cgagatccgt	7200
gcctcgtgga cgccgcgtga gggctcgggc ggggttcgg acacggccgc ccatctggtc	7260
gcgtggtccg atctgctggc gcaggtcgcg gggctgccgc cggccgctcc gggggacgcg	7320

tccgtggtga cgctgccga gcggagggg ccgcagtcgc gctgagcctc ccgctgcggg	7380
gtcccgacga ccggcctcct gcatctctct ttgtcgatac ggccactttc ggaagcgttg	7440
cgacgcagac cgaaaccgtt cgatcttcga atgacgatac gtgcgtccga attgcccgga	7500
ttgttactca tcacttcgtg atcattcgtt aaaggacacc aggtttgctg ccgaagacga	7560
ctgtgacctt gaaagcacgg ttctgccgc cttaccccc acgagccggc ccgtcccgca	7620
ccccaggagg cctggtgttt aattaatgac ccttcgctc cggcccgat ccgcccgtcc	7680
ccgatccgcc cggccacggc caagcgaaca cgttctttca ctctctgac cggaatacga	7740
cccaccggcg ccggtcacgg agcacaaccg tcgacggcg ccttcgcggc acggataccc	7800
ttgacaggtg accgacgcc acgacaccgc agcagacagt tcaactgcga ccaccggagg	7860
cgctcctccg gacgacggcg gatcttctgt tacggaggcg ccgacccct tgctggaacc	7920
ccgcgagggc attccgccg tgatagcgga cgaggccgcc ctcccgagg cggtcgccgc	7980
cttcgcggcc ggcagcggac ccgtcgccgt ggacgccgag cgcgcctccg ggtaccgcta	8040
cggccagcgc gcctacctg tccagctgcg ccgcgagggt gcgggtaccg cgtgatcga	8100
ccccgtggcc tgccccgacc tgtccgccct cggcgaggcg ctgtccggcg tcgagtgggt	8160
gctgcacgcc gccaccagg acctgccctg tctgcgcgag ataggcatgg tgccctccc	8220
cctcttcgac accgagctgg ccggccgcct tgccgggttc ccccgctcg ggctcggcg	8280
gatggtcgag aacgtgctcg gcttcgtcct ggagaagggc cactccgccg tcgactggtc	8340
caccgtccg ctgccgagc cctggctgcg gtacgccgcc ctgcacgtcg aactgctggt	8400
cgatctgcgg gacgccctgg agaaggagct ggaccgccag ggcaagctgg actgggccc	8460
gcaggagttc gacgcgatcg cctcggcccc gccgccggag ccccgcaagg acccctggcg	8520
ccgcacctcc ggcatgcaca aggtgcgccg gcgccgccag atggcggttg tcggggagct	8580
gtgggagacc cgcgaccgga tcgccggcg ccgtgacgtc tccccggca aggtgctttc	8640
cgacgcggcg atcgtggagg ccgcgctcg gctgccgcc aacctgcacg ccatggccgc	8700
gctcaacggg ttccgggcagc gggtagggcg gcgccagctg gagcagtggc aggcggccgt	8760

cgaccgcgcg aaggcgtga gcgaggccca gctgccgcag cccggccagc cggtgaccgg	8820
ccctccgccg ccgcgcgcct gggcggacaa ggaccccggt gccgcggccc ggctgtcggc	8880
ggcccgcgcg ggggtcgccg aactcgccga gcggctgaac atgccgccgg agaacctgat	8940
caccccggaac acggtgcgca gggctctgct ggagccgccg gggcccgacg agcgggtccgt	9000
cgcgcggcg ctgacggcgc acggggcacg cgcgtggcag gtcgaccagg tcaactccgt	9060
gctggtggcc gcgctggcta ctctgtgcc ccccgcatga acccgttgta gatgaagcgc	9120
tggaggaaca ggaagacgat caaggtgggc aggatgacca ggaccgcgcc cgccgagatc	9180
gtctcccagt gcgcgccgaa ggggcccttg aagcggaaaca gggacgtcga gatgaccccc	9240
aggtcctcgg agggcatgta gaggaagggg atgtagaagt cgttgtagac gttgatcccc	9300
tttacgatca ccaccgtgc gatcgccggc ttgagcagcg ggaagatcac ctgtcggtag	9360
acggtgaacg cgttggcgcc gtccaggcgc gccgcctcgt ccaggagac ggggatggag	9420
cggatgaact gcaggaagac gtagatcgag acgatgtccg tgcccatgta gagggcgatc	9480
ggcgcccaca ggctgtcgaa catgccgaag ctgttgacga tctggaaggt cgccacctgg	9540
gtggtcacc cggggaccag cgcggccagc aggaacagcg ccacgaccag ctagcgaatt	9600
cctgcaggtc cccggggatc ggtcttgctt tgctcgtcgg tgatgtactt caccagctcc	9660
gcgaagtcgc tcttcttgat ggagcgcgat gggacgtgct tggcaatcac gcgcaccccc	9720
cggccgtttt agcggctaaa aaagtcattg ctctgccctc gggcggacca cgcccatcat	9780
gaccttgcca agctcgtcct gttctcttc gatcttcgcc agcagggcga ggatcgtggc	9840
atcaccgaac cgcgccgtgc gcgggtcgtc ggtgagccag agtttcagca ggccgccag	9900
goggcccagg tcgccattga tgcgggccag ctgcgggacg tgctcatagt ccacgacgcc	9960
cgtgattttg tagccctggc cgacggccag caggtaggcc gacaggctca tgccggccgc	10020
cgccgccttt tcctcaatcg ctcttcgttc gtctggaagg cagtacacct tgataggtgg	10080
gctgcccttc ctggttggct tggtttcac agccatccgc ttgccctcat ctgttacgcc	10140

ggcggtagcc ggccagcctc gcagagcagg attcccgttg agcaccgcca ggtgcgaata	10200
agggacagtg aagaaggaac acccgctcgc ggggtgggcct acttcaccta tcctgcccgg	10260
ctgacgccgt tggatacacc aaggaaagtc tacacgaacc ctttggcaaa atoctgtata	10320
tcgtgcgaaa aaggatggat ataccgaaaa aatcgctata atgaccccgga agcagggtta	10380
tgcagcggaa aagatccgtc gacctgcagc ccgggggata cccgggtacc gagctcgggc	10440
tggggctcgg ggccgccggt gagctggtag acgaaggcgc ccgagtctcc ttcgttcact	10500
gcgtgccact cgtggtgcgg gtacttccgg cgcaactgc tgtcgtccat gggcggcatc	10560
atggcagagg cggagacgcc gttccgcgcc tttcgtcggg gcccgtaggg tttcggacat	10620
tcttgtgcgg ggtggggggg cgccggcgga cccggtgcgc ccggcgtcgc	10670

<210> 9
 <211> 40
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 9	
agcacagcta gcttctagaa gcttcattca aaggccggca	40

<210> 10
 <211> 43
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 10	
gccagtgaat tctgcagctc gagcagagca ggattcccggt tga	43

<210> 11
 <211> 7173
 <212> ADN
 <213> Secuencia Artificial

<220>

<223> PI 痘 mido Sint 驕 ico pKS-T-orf18pfrd

<400> 11

gtggcacttt tgggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt	60
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa	120
ggaagagtat gagtattcaa catttccgtg tgcgccttat tccctttttt gcggcatttt	180
gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt	240
tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt	300
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg	360
tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga	420
atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa	480
gagaattatg cagtgtctgcc ataaccatga gtgataacac tgcggccaac ttacttctga	540
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa	600
ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca	660
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaaactggc gaactactta	720
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac	780
ttctgcgctc ggcccttcg gctggctggt ttattgctga taaatctgga gccggtgagc	840
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag	900
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atogctgaga	960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt	1020
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata	1080
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag	1140
aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa	1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt	1260

ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc	1320
cgtagttagg ccaccacttc aagaactctg tagcaccgoc tacatacctc gctctgctaa	1380
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa	1440
gacgatagtt accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc	1500
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa	1560
gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggctcgaa	1620
caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg	1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctctcaggg gggcggagcc	1740
tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggcccttttc tggccttttg	1800
ctcacatgtt ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg	1860
agtgagctga taccgctcgc gcgagccgaa cgaccgagcg cagcgagtca gtgagcgagg	1920
aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgccg attcattaat	1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg	2040
tgagttagct cactcattag gcaccccagg ctttacactt tatgcttcg gctcgtatgt	2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg	2160
ccaagcgcgc aattaaccct cactaaaggg aacaaaagct gggtaaccgg cccccctcg	2220
agcagagcag gattcccgtt gagcaccgcc aggtgcgaat aaggacagt gaagaaggaa	2280
caccgctcg cgggtgggcc tacttcacct atcctgcccg gctgacgcc ttggatacac	2340
caaggaaagt ctacacgaac cttttggcaa aatcctgtat atcgtgcgaa aaaggatgga	2400
tataccgaaa aaatcgctat aatgaccccg aagcagggtt atgcagcgga aaatgcagct	2460
cacggtaact gatgccgtat ttgcagtacc agcgtacggc ccacagaatg atgtcacgct	2520
gaaaatgccg gcctttgaat gaagcttgcc ggggccgacg cggcggcgcg cctcgtccat	2580
cgggacgcgc cagtcgacgc cgacgatgtc cgcgccggcc tcgcccata gcccagcag	2640
ctcgccggtg ccgacgccga agtggatgcg cgggacgccg tggccggcca ccgcgcggaa	2700

caccttcgcc gaggcgggca gcaccgaacg ccggtagtcg gagggggcga gcgcgccggc	2760
ccaggagtcg aagagctgca cggccgaggc gccggcccgg atctggacgt cgaggaaggc	2820
cgccgtgatg tcggcgaggc ggtcgagcag gtcggcccag agctcggggc cgccgtacat	2880
catgccttg gcgttctcgt acgtacggga cgggccgccc tcgacgaggt aactcgcaag	2940
ggtgaacggg gcgcccgcga aaccgatcag cggggtggac ccgagctcac gggtcagcat	3000
gccgatggcc tcggtgacgt aggagacgtc ctccggggtc aggtcgcgca gccgggcgag	3060
gtcggcgcgg gtgcgcaccg gctgctcgac gaccggggcg atgccgggct tgatgtcgag	3120
gtcgatgccg atggccttga gcgggacgac gatgtcgtg aagtagatcg ccgcgtcgac	3180
gtggtggcgg cgcaccggct ggagggtgat ctcggtgacc agctcgggcc gcatgcagga	3240
gtcgagcatc ccgatgccct cgcgcacctt gcggtactcc ggcagtgagc gcccggcctg	3300
ccgcatgaac cacaccggcg tgtgcggcac cggctcgcgc cggcacgcct tgaggaaggc	3360
tgattcccgg acggcgtcgt tcttgacggg gtcgggggtc gcggtaggcg tctggccctt	3420
cgggctcgtg ttggcactca caccggccag tctcgcacga cccgtccgcc gccccggacc	3480
ccaggggcgt ccggcgggcg gggccgcggc gcaggatccg gacagcgccg ggcggcgcgg	3540
gtgtccctcc ctgcgccggg ggcccgttcc gcttaatctt cccggcatgg ctgcggctca	3600
gggacgactg tcggacggcg ctggcggaat ggacgaaccg aaggagggcg ggggggatcc	3660
cgggcacgga ggtgcgcctc cgccgccctt ccgggccgct gtcgaggcgc tgcagagcgc	3720
ccggctgcgg ccgcagatcg aggtggagac ggtgcccgcg ccgaaacggc tcgccccgta	3780
cgcgcacgcg ctggaggcgg cggtcgtcga cggcgaggag gatctggccg acggccggct	3840
ggtgctgctg tgcgaccogg ccggacacga cgcctggcgg gggaccttcc gtctggtgac	3900
gctggtgcgc gccgagctgg agccggagat ggcggcggat ccgctgctgc cggacgtgtg	3960
ctggtcctgg ctgaccggcg cgttggcggc gcgcggcctg tcgtacggcg agccgagcgg	4020
cacggtgacg cgggcgagtt cgcactactt cggcgggctg tccgcgcggc ccgccgcctc	4080

ccagatcgag atccgtgcct cgtggacgcc gcgtgagggt ctgggcgggg ttccggacac	4140
ggccgcccac ctggtcgcgt ggtccgatct gctggcgag gtcgcggggc tgccgccggc	4200
cgctccgggg gacgcgtccg tggtagcgt gccgcagcgg agggggccgc agtcgcgtg	4260
agcctcccgc tgcggggtcc cgacgaccgg cctcctgcat ctctctttgt cgatacggcc	4320
actttcgga gcttgcgac gcagaccgaa accgttcgat cttcgaatga tcgatcgtgc	4380
gtccgaattg cccggttgt tactcatcac ttctgatca ttcttaaag gacaccaggt	4440
ttgtgccga agacactgt gacctgaaa gcacggttcg tccgccttc acccccacga	4500
gccggcccgt cccgcacccc aggaggcctg gtgtttaatt aatgaccctt ccgtccgcc	4560
ccgatccgc ccgtcccga tccgccgcc caccggcaag cgaacacgtt ctttactct	4620
tctgaccgga atacgacca ccggcgccc tcacggagca caaccgtcg cggcgccctt	4680
cgcggcacgg ataccctga caggtgaccg acgcccacga caccgcagca gacagttcac	4740
tgcgcaccac cggaggcgt cctccggacg acggcggatc ttctgttacg gaggcgccga	4800
cccccttgt ggaacccgc gagggcattc cgccgtgat agcggacgag gccgccctcg	4860
ccgaggcgt cgccgccttc gggccggca gcggaccgt cgccgtggac gccgagcgcg	4920
cctccgggta ccgctacggc cagcgccct acctcgtcca gtcgcgccgc gaggtgcgg	4980
gtaccgcgt gatcgacccc gtggcctgcc ccgacctgc cgccctggc gaggcgtgt	5040
ccggcgtcg gtgggtgctg caccgcgca ccaggacct gccctgtctg cgcgagatag	5100
gcatggtgcc ctccgcctc ttcgacaccg agctggccgg ccgccttgcc gggttcccc	5160
gcgtcgggct cggcgcatg gtcgagaacg tgctcggctt cgtcctggag aaggccact	5220
ccgcgtcg ctggtccacc cgtccgtgc ccgagccctg gtcgcgtac gccgccctcg	5280
acgtcgaact gctggtcgat ctgcgggacg ccctggagaa ggagctggac gccagggca	5340
agctggactg ggcccgagc gattcgacg cgatcgctc ggccccgcg ccgagcccc	5400
gcaaggaccc ctggcgccgc acctccggca tgcacaagg gcgcggcgcc gccagatgg	5460
cgggtgtgcg ggagctgtg gagaccgcg accggatcgc ccggcgccgt gacgtctccc	5520

cggcaaggt gctttccgac gcggcgatcg tggaggccgc gctcgcgtg cccccaacc	5580
tgcacgcat ggccgcgctc aacgggttcg ggcagcgggt gggcgggcgc cagctggagc	5640
agtggcaggc ggccgtcgac cgcgcgaagg cgctgagcga ggcccagctg ccgcagccc	5700
gccagccggt gaccggccct ccgccgccgc gcgcctgggc ggacaaggac ccggttgccg	5760
cggccgggt gtcggcgcc cgcgcggggg tcgccgaact cgccgagcgg ctgaacatgc	5820
cgccggagaa cctgatcacc ccggacacgg tgcgcagggt ctgctgggag ccgccggggc	5880
ccgacgagcg gtccgtcgcc gcggcgctga cggcgcacgg ggcacgcgcg tggcaggtcg	5940
accaggtcac tccgtgctg gtggccgcgc tggctacttc gtcgcccccc gcatgaaccc	6000
gttgtagatg aagcgtgga ggaacaggaa gacgatcaag gtgggcagga tgaccaggac	6060
cgcgccgcc gagatcgtct ccagtgccgc gccgaagggg ccctgaagc ggaacaggga	6120
cgtcgagatg acccccaggt cctcgagggg catgtagagg aaggggatgt agaagtcgtt	6180
gtagacgttg atccccitta cgatcaccac cgtcgcgac gccggcttga gcagcgggaa	6240
gatcaccttg cggtagacgg tgaacgcgtt ggccgcgtcc aggcgcgcgc cctcgtccag	6300
ggagacgggg atggagcgga tgaactgcag gaagacgtag atcgagacga tgtccgtgcc	6360
catgtagagg gcgatcggcg ccacacaggct gtcgaacatg ccgaagctgt tgacgatctg	6420
gaaggtcgcc acctgggttg tcaccccggg gaccagcgcg gccagcagga acagcggcac	6480
gaccagctag ttctagagcg gccgccaccg cggtagagct ccaattcgcc ctatagttag	6540
togtattacg cgcgctcact ggccgtcgtt ttacaacgtc gtgactggga aaaccctggc	6600
gttaccacaac ttaatgcct tgcagcacat cccctttcg ccagctggcg taatagcgaa	6660
gaggcccgca ccgatcgccc ttccaacag ttgcgcagoc tgaatggcga atggaaattg	6720
taagcgttaa tattttgtta aaattcgct taaattttg ttaaatcagc tcatttttta	6780
accaataggc cgaaatcggc aaaatccctt ataatcaaa agaatagacc gagatagggt	6840
tgagtgttgt tccagtttg aacaagagtc cactattaaa gaacgtggac tccaacgtca	6900

aagggcgaaa aaccgtctat cagggcgatg gccactacg tgaacctca ccctaataca	6960
gttttttggg gtcgaggtgc cgtaaagcac taaatcgaa ccctaaagg agccccgat	7020
ttagagcttg acggggaaag ccggcgaac tggcgagaaa ggaagggaag aaagcgaaag	7080
gagcgggcgc tagggcgctg gcaagtgtag cggtcacgt gcgcgtaacc accacacccg	7140
ccgcgcttaa tgcgccgcta cagggcgctg cag	7173

<210> 12
 <211> 29
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Gebador de Oligonucle · ido Sint 騁 ico

<400> 12	
gaatggccat ggttcatgtg cagctccat	29

<210> 13
 <211> 30
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Gebador de Oligonucle · ido Sint 騁 ico

<400> 13	
tctcgaggat ccgaatagga acttcggaat	30

<210> 14
 <211> 6034
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Pl 痘 mido Sint 騁 ico pKS-T-orf18pfrd-AmR

<400> 14	
gtggcacttt tcggggaaat gtgcgcggaa ccctatttg tttatttttc taaatacatt	60

caaatatgta tccgctcatg agacaataac cctgataaat gttcaataa tattgaaaaa	120
--	-----

ggaagagtat gagtattcaa catttccgtg togcccttat tccctttttt gcggcatttt	180
gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt	240
tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt	300
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg	360
tattatcccg tattgaacgc gggcaagagc aactcggctg ccgcatacac tattctcaga	420
atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa	480
gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga	540
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa	600
ctcgccctga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca	660
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaaactggc gaactactta	720
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac	780
ttctgcgctc ggcccttcg gctggctggg ttattgtgta taaatctgga gccggtgagc	840
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag	900
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga	960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt	1020
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata	1080
atctcatgac caaaatccct taacgtgagt ttctgttcca ctgagcgtca gaccccgtag	1140
aaaagatcaa aggatcttct tgagatcctt ttttctgog cgtaatctgc tgcttgcaaa	1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt	1260
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc	1320
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa	1380
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa	1440
gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc	1500

ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa	1560
gcgccacgct tcccgaaggg agaaaggcgg acaggtatoc ggtaagcggc agggtcggaa	1620
caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg	1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc	1740
tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg	1800
ctcacatgtt ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg	1860
agtgagctga taccgctogc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg	1920
aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgccg attcattaat	1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg	2040
tgagttagct cactcattag gcaccccagg ctttacactt tatgcttcg gctcgtatgt	2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg	2160
ccaagcgcgc aattaaccct cactaaaggg aacaaaagct gggtaaccggg cccccctcg	2220
agcagagcag gattcccgtt gagcaccgcc aggtgcgaat aaggacagt gaagaaggaa	2280
caccgctcg cgggtgggcc tacttcacct atcctgcccg gctgacgccg ttggatacac	2340
caaggaaagt ctacacgaac cttttggcaa aatcctgtat atcgtgcgaa aaaggatgga	2400
tataccgaaa aaatcgctat aatgaccccg aagcagggtt atgcagcgga aaatgcagct	2460
cacggtaact gatgccgtat ttgcagtacc agcgtacggc ccacagaatg atgtcacgct	2520
gaaaatgccg gcctttgaat gaagcttgcc ggggccgacg cggcggcgcg cctcgtccat	2580
cgggacgcgc cagtcgacgc cgacgatgtc cgcgccggcc tcgcccata gcccagcag	2640
ctcgccggtg ccgacgccga agtggatgcg cgggacgccg tggccggcca ccgcgcggaa	2700
caccttcgcc gaggcgggca gcaccgaac ccggtagtgc gagggggcga gcgcgccggc	2760
ccaggagtcg aagagctgca cggccgaggc gccggcccg atctggacgt cgaggaaggc	2820
cgccgtgatg tcggcgaggc ggtcgagcag gtcggccag agctcggggt cgccgtacat	2880
catcgccttg gcgttctcgt acgtacggga cgggccgccc tcgacgaggt aactcgcaag	2940

ggtgaacggg ggcgccgga aaccgatcag cggggtggac ccgagctcac gggtcagcat	3000
gccgatggcc tcggtgacgt aggagacgtc ctccggggtc aggtcgcgca gccgggag	3060
gtcggcgcggt gtgcgcaccg gctgctcgac gaccggggcc atgccgggct tgatgtcgag	3120
gtcgatgccg atggccttga gggggacgac gatgtcgctg aagtagatcg ccgcgtcgac	3180
gtggtggcgg cgcaccggct ggagggtgat ctcggtgacc agctcgggcc gcatgcagga	3240
gtcgagcatc ccgatgccct cgcgcacctt gcggtactcc ggcatgagc gcccggcctg	3300
ccgcatgaac cacaccggcg tgtgcggcac cggctcgcgc cggcacgcct tgaggaaggc	3360
tgattcccg acggcgctgt tcttgacggg gtcgggggtc gcggtaggcg tctggccctt	3420
cgggctcgtg ttggcactca caccggccag tctcgcacga cccgtccgcc gccccggacc	3480
ccaggggcgt ccggcgggcg gggccgcggt gcaggatccg aataggaact tcggaatagg	3540
aacttcatga gctcagccaa tcgactggcg agcggcatcg cattcttcgc atccgcctc	3600
tggcggatgc aggaagatca acggatctcg gccagttga ccaggggctg tgcacacaat	3660
gtcgcgggag cggatcaacc gagcaaaggc atgaccgact ggaccttcct tctgaaggct	3720
cttctccttg agccacctgt ccgccaaggc aaagcgctca cagcagtggc cattctcgag	3780
ataatcgacg cgtaccaact tgccatcctg aagaatggtg cagtgtctcg gcacccata	3840
gggaaccttt gccatcaact cggcaagatg cagcgtcgtg ttggcatcgt gtcccacgcc	3900
gaggagaagt acctgcccat cgagttcatg gacacgggag accgggcttg caggcgagt	3960
aggtggcagg ggcaatggat cagagatgat ctgctctgcc tgtggccccg ctgccgaaa	4020
ggcaaatgga tgggcgctgc gctttacatt tggcaggcgc cagaatgtgt cagagacaac	4080
tccaaggtcc ggtgtaacgg gcgacgtggc aggatcgaac ggctcgtcgt ccagacctga	4140
ccacgagggc atgacgagcg tccctcccg acccagcgca gcacgcagg cctcgatcag	4200
tccaagtggc ccatcttcga ggggccggac gctacggaag gagctgtgga ccagcagcac	4260
accgccgggg gtaaccccaa ggttgagaag ctgaccgatg agctcggctt ttgccattc	4320

gtattgcacg acattgcact ccaccgctga tgacatcagt cgatcatagc acgatcaacg	4380
gcactgttgc aaatagtcgg tgggtgataaa ottatcatoc ctttttgctg atggagctgc	4440
acatgaacca tggccgcgct caacgggttc gggcagcggg tggggcggcg ccagctggag	4500
cagtggcagg cggccgtcga ccgcgcgaag gcgctgagcg aggcccagct gccgcagccc	4560
ggccagccgg tgaccggccc tccgcgcgcg cgcgcctggg cggacaagga ccccgttgcc	4620
gcgcccggc tgtcggcggc ccgcgcgggg gtgcgcgaac tcgccgagcg gctgaacatg	4680
ccgccggaga acctgatcac cccggacacg gtgcgcaggg tctgctggga gccgccgggg	4740
cccgcagagc ggtccgtcgc cgcggcgctg acggcgcacg gggcacgcgc gtggcaggtc	4800
gaccaggtca ctcccgctgt ggtggcgcg ctggctactt cgtcgcccc cgcatgaacc	4860
cgttgtagat gaagcgtgg aggaacagga agacgatcaa ggtgggcagg atgaccagga	4920
ccgcgcccgc cgagatcgtc tcccagtgcg cgccgaaggg gcccttgaag cggaacaggg	4980
acgtcgagat gacccccagg tcctcggagg gcatgtagag gaaggggatg tagaagtcgt	5040
tgtagacgtt gatccccctt acgatcacca ccgtcgcgat cgccggcttg agcagcggga	5100
agatcacctt gcggtagacg gtgaacgcgt tggcgccgtc caggcgcgcc gcctcgtcca	5160
gggagacggg gatggagcgg atgaactgca ggaagacgta gatcgagacg atgtccgtgc	5220
ccatgtagag ggcgatcggc gccacaggc tgtcgaacat gccgaagctg ttgacgatct	5280
ggaaggtcgc cacctgggtg gtcaccccg ggaccagcgc gcccagcagg aacagcgcca	5340
cgaccagcta gttctagagc ggccgccacc gcggtggagc tccaattcgc cctatagtga	5400
gtcgtattac gcgcgctcac tggcgcgtgt tttaaacgt cgtgactggg aaaaccctgg	5460
cgttacccaa cttaatcgcc ttgcagcaca tcccccttc gccagctggc gtaatagcga	5520
agaggcccg accgatcgcc ctcccaaca gttgcgcagc ctgaatggcg aatggaaatt	5580
gtaagcgta atattttgtt aaaattcgcg ttaaattttt gttaaatacag ctcatTTTTT	5640
aaccaatagg ccgaaatcgg caaaatccct tataaatcaa aagaatagac cgagataggg	5700
ttgagtgttg ttccagtttg gaacaagagt ccactattaa agaacgtgga ctccaacgtc	5760

aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc accctaataca	5820
agtttttttg ggtcgagggtg ccgtaaagca ctaaatacgga accctaaagg gagccccga	5880
tttagagctt gacggggaaa gccggcgaac gtggcgagaa aggaaggga gaaagcgaaa	5940
ggagcgggcg ctagggcgct ggcaagtgtg gcggtcacgc tgcgcgtaac caccacacc	6000
gccgcgctta atgcgcgct acagggcgcg tcag	6034

<210> 15
 <211> 40
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 15	
agcacagcta gcttctagaa gcttcattca aaggccggca	40

<210> 16
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 16	
aggcagctcg agcatatgac tagtcagagc aggattcccg ttga	44

<210> 17
 <211> 30
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 17	
gaatggcata tggttcatgt gcagctccat	30

<210> 18
<211> 30
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 18
tctagaacta gtgaatagga acttcggaat 30

<210> 19
<211> 8113
<212> ADN
<213> Secuencia Artificial

<220>
<223> PI 痘 mido Sint 騁 ico pKS-orf18ifd-T-AmR(NS)

<400> 19
gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 60
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa 120
ggaagagtat gagtattcaa catttcctgt tcgcccttat tccctttttt gcggcatttt 180
gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
tggttgcaag agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt 300
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg 360
tattatcccg tattgaagcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420
atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 540
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attactggc gaactactta 720
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780

ttctgcgctc ggcccttcgc gctggctggt ttattgctga taaatctgga gccggtgagc	840
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag	900
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga	960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt	1020
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata	1080
atctcatgac caaaatccct taacgtgagt tticgttcca ctgagcgtca gaccccgtag	1140
aaaagatcaa aggatcttct tgagatcctt tttttctgog cgtaatctgc tgccttgcaaa	1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt	1260
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc	1320
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa	1380
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa	1440
gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc	1500
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa	1560
gcgccacgt tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa	1620
caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg	1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc	1740
tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg	1800
ctcacatgtt ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg	1860
agtgagctga taccgctcgc gcgagccgaa cgaccgagcg cagcgagtca gtgagcgagg	1920
aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgccg attcattaat	1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg	2040
tgagttagct cactcattag gcaccccagg ctttacactt tatgcttcg gctcgtatgt	2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg	2160

ccaagcgcgc aattaacct cactaaaggg aacaaaagct gggtaccggg cccccctcg	2220
agcatatgtt catgtgcagc tccatcagca aaaggggatg ataagtttat caccaccgac	2280
tatttgcaac agtgccgttg atcgtgctat gatcgactga tgtcatcagc ggtggagtgc	2340
aatgtcgtgc aatacgaatg gcgaaaagcc gagctcatcg gtcagcttct caaccttggg	2400
gttaccctcg gcggtgtgct gctgggtccac agctccttcc gtagcgtccg gccctcga	2460
gatgggccac ttggactgat cgaggccctg cgtgctgcgc tgggtccggg agggacgctc	2520
gtcatgccct cgtggtcagg tctggacgac gagccgttcg atcctgccac gtgcccgtt	2580
acaccggacc ttggagtgtt ctctgacaca ttctggcgcc tgccaaatgt aaagcgcagc	2640
gcccattcat ttgcctttgc ggacaggggg ccacaggcag agcagatcat ctctgatcca	2700
ttgccctgc caccctactc gcctgcaagc ccggtcgccc gtgtccatga actcgatggg	2760
caggtacttc tcctcggcgt gggacacgat gccaacacga cgctgcatct tgccgagttg	2820
atggcaaagg ttccctatgg ggtgccgaga cactgcacca ttcttcagga tggcaagttg	2880
gtacgcgtcg attatctoga gaatgaccac tgctgtgagc gctttgcctt ggoggacagg	2940
tggctcaagg agaagagcct tcagaaggaa ggtccagtcg gtcatgcctt tgctcggttg	3000
atccgctccc gcgacattgt ggcgacagcc ctgggtcaac tgggccgaga tccgttgatc	3060
ttctgcac cgcagaggc gggatgcgaa gaatgcgatg ccgctcgcca gtcgattggc	3120
tgagctcatg aagttcctat tccgaagttc ctattcacta gtcagagcag gattcccgtt	3180
gagcaccgcc aggtgcgaat aaggacagc gaagaaggaa caccgctcg cgggtgggcc	3240
tacttcacct atcctgccc gctgacgcc ttggatacac caaggaaagt ctacacgaac	3300
cctttggcaa aatcctgtat atcgtgcgaa aaaggatgga tataccgaaa aaatcgctat	3360
aatgaccccg aagcagggtt atgcagcgga aaatgcagct cacggtaact gatgccgtat	3420
ttgcagtacc agcgtacggc ccacagaatg atgtcacgct gaaaatgccg gcctttgaat	3480
gaagcttgcc ggggccgacg cggcgggcgg cctcgtccat cgggacgcgc cagtcgacgc	3540
cgacgatgtc cgcgccggcc tcgcccata gccccagcag ctgcgccgtg ccgacgccga	3600

agtggatgcg cgggacgccg tggccggcca ccgcgcggaa caccttcgcc gaggcgggca	3660
gcaccgaacg ccggtagtcg gagggggcga gcgcgccggc ccaggagtcg aagagctgca	3720
cggccgaggc gccggccccg atctggacgt cgaggaaggc cgccgtgatg tcggcgaggc	3780
ggtcgagcag gtcggcccag agctcggggt cgccgtacat catcgccctg gcgttctcgt	3840
acgtacggga cgggcgcgcc tcgacgaggt aactcgcaag ggtgaacggg gcgccccgca	3900
aaccgatcag cggggtggac ccgagctcac gggtcagcat gccgatggcc tcggtgacgt	3960
aggagacgtc ctccggggtc aggtcgcgca gccgggcgag gtcggcgcgg gtgcgcaccg	4020
gttgctcgac gaccggggcc atgccgggct tgatgtcgag gtcgatgccg atggccttga	4080
gcgggacgac gatgtcgtg aagtagatcg ccgcgtcgac gtggtggcgg cgcaccggct	4140
ggagggtgat ctcggtgacc agctcgggcc gcatgcagga gtcgagcatc ccgatgccct	4200
cgcgcacctt gcggtactcc ggcagtgagc gcccggcctg ccgcatgaac cacaccggcg	4260
tgtgcggcac cggctcgcgc cggcacgcct tgaggaaggc tgattcccgg acggcgtcgt	4320
tcttgacggg gtcgggggtc gcggtaggcg tctggccctt cgggctcgtg ttggcactca	4380
caccggccag tctcgacga cccgtccgcc gccccggacc ccaggggcgt ccggcgggcg	4440
gggcgcgggt gcaggatccg gacagcgccg ggcggcgcgg gtgtccctcc ctgcgccggg	4500
ggcccgttcc gcttaatctt cccggcatgg ctgcggctca gggacgactg tcggacggcg	4560
ctggcggaat ggacgaaccg aaggagggcg ggggggatcc cgggcacgga ggtgcgcctc	4620
cgccgccctt ccgggcccgt gtcgaggcgc tgcagagcgc ccggtgcgg ccgcagatcg	4680
aggtggagac ggtgcccgcg ccgaaacggc tcgccccgta cggcacgcg ctggaggcgg	4740
cggctgtcga cggcgaggag gatctggccg acggccggct ggtgctgctg tgogaccgg	4800
ccggacacga cgcctggcgg gggaccttcc gtctggtgac gctggtgcgc gccagctgg	4860
agccggagat ggcggcggat ccgtgctgc cggacgtgtg ctggtcctgg ctgaccggcg	4920
cgctggcggc gcgcggcctg tcgtacggcg agccgagcgg cacggtgacg cgggcgagtt	4980

cgcactactt cggcgggctg tccgcgcggc ccgccgcctc ccagatcgag atccgtgcct	5040
cgtggacgcc gcgtgagggt ctgggcgggg ttccggacac ggccgoccat ctggtcgcgt	5100
ggtccgatct gctggcgcag gtgcgggggc tgccgccggc cgctccgggg gacgcgtccg	5160
tggtgacgct gccgcagcgg agggggccgc agtcgcgctg agcctcccgc tgcggggtcc	5220
cgacgaccgg cctcctgcat ctctctttgt cgatacggcc actttcggaa gcgttgcgac	5280
gcagaccgaa accgttcgat cttcgaatga tcgatcgtgc gtccgaattg cccggattgt	5340
tactcatcac ttctgatca ttcgttaaag gacaccaggt ttgctgccga agacgactgt	5400
gaccttgaag gcacggttcg tccgccttc acccccacga gccggcccgt cccgcacccc	5460
aggaggcctg gtgtttaatt aatgaccctt ccgtcccgc cccgatccgc ccgtcccga	5520
tccgcccgcc cacggccaag cgaacacgtt ctttactct tctgaccgga atacgacca	5580
ccggcgcccg tcacggagca caaccgtcga cgggcgcctt cgcggcacgg atacccttga	5640
caggtgaccg acgccacga caccgcagca gacagttcac tgcgcaccac cggaggcgct	5700
cctccggacg acggcggatc ttctgttacg gaggcgccga cccccttgct ggaacccgc	5760
gagggcattc cgcctgtgat agcggacgag gccgccctcg ccgaggcggt cgcgccttc	5820
gcggccggca gcggaccgt cgcctggac gccgagcgcg cctccgggta ccgctacggc	5880
cagcgcgcct acctcgtcca gctgcgccgc gaggtgcgg gtaccgcgt gatcgacccc	5940
gtggcctgcc ccgacctgtc cgccctcggc gaggcgtgt ccggcgtcga gtgggtgctg	6000
cacgccgcca ccaggacct gccctgtctg cgcgagatag gcatggtgcc ctcccgcctc	6060
ttcgacaccg agctggccgg ccgccttgcc gggttcccc gcgtcgggct cggcgcgatg	6120
gtcgagaacg tgctcggctt cgtcctggag aagggccact ccgccgtcga ctggtccacc	6180
cgtccgctgc ccgagccctg gctgcggtac gccgccctcg acgtcgaact gctggtcgat	6240
ctgcgggacg ccctggagaa ggagctggac gcccagggca agctggactg ggcccggcag	6300
gagttcgacg cgatcgctc ggccccgccg ccggagcccc gcaaggaccc ctggcgccgc	6360
acctccggca tgcacaaggt gcgccggcgc gccagatgg cggtggtgcg ggagctgtgg	6420

gagacccgcg accggatogc ccggcgccgt gacgtctccc ccggcaaggt gctttccgac	6480
gcggcgatcg tggaggccgc gctcgcgctg cccgccaacc tgcacgccat ggccgcgctc	6540
aacgggttcg ggccagcgggt ggggcggcgc cagctggagc agtggcaggc ggccgtcgac	6600
cgcgcgaagg cgctgagcga gggccagctg ccgcagcccg gccagccggt gaccggccct	6660
ccgcccgcgc ggccttgggc ggacaaggac cccgttgccg cggcccggct gtccggcgcc	6720
cgcgcggggg tcgccgaact cgcgcgagcg ctgaacatgc cgcggagaa cctgacacc	6780
ccggacacgg tgcgcagggt ctgctgggag ccgcccgggc ccgacgagcg gtccgtcgcc	6840
gcggcgctga cggcgacagg ggcacgcgcg tggcaggctg accaggtcac tcccgtgctg	6900
gtggccgcgc tggctacttc gtccgcccc gcataaacc gttgtagatg aagcgctgga	6960
ggaacaggaa gacgatcaag gtgggcagga tgaccaggac cgcgcccgc gagatcgtct	7020
cccagtgcgc gccgaagggg cccttgaagc ggaacaggga cgtcgagatg acccccaggt	7080
cctcggaggg catgtagagg aaggggatgt agaagtcgtt gtagacgttg atccccctta	7140
cgatcaccac cgtcgcgac gccggcttga gcagcgggaa gatcaccttg cggtagacgg	7200
tgaacgcgtt ggccgctcc aggcgcgccg cctcgtccag ggagacgggg atggagcgga	7260
tgaactgcag gaagacgtag atcgagacga tgtccgtgcc catgtagagg gcgatcgcg	7320
cccacaggct gtcgaacatg ccgaagctgt tgacgatctg gaaggtcgcc acctgggtgg	7380
tcacccggg gaccagcgc gccagcagga acagcgccac gaccagctag ttctagagcg	7440
gccgccaccg cgggtggagct ccaattcgcc ctatagttag tcgtattacg cgogctcact	7500
ggccgtcgtt ttacaacgtc gtgactggga aaacctggc gttaccaac ttaatgcct	7560
tgcagcacat cccctttcg ccagctggcg taatagcga gagggccgca ccgatcgccc	7620
ttccaacag ttgcgcagcc tgaatggcga atggaaattg taagcgttaa tatittgta	7680
aaattcgcgt taaatttttg ttaaatcagc tcatttttta accaataggc cgaaatcggc	7740
aaaatccctt ataaatcaaa agaatagacc gagatagggt tgagtgttgt tccagtttgg	7800

aacaagagtc cactattaaa gaacgtggac tccaacgtca aagggcgaaa aaccgtctat	7860
caggcgatg gccactacg tgaaccatca ccctaataa gttttttggg gtogaggtgc	7920
cgtaaagcac taaatcggaa ccctaaaggg agccccgat ttagagcttg acggggaaaag	7980
ccggcgaacg tggcgagaaa ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg	8040
gcaagtgtag cggtcacgct gcgcgtaacc accacaccg ccgcgcttaa tgcgccgcta	8100
cagggcgct cag	8113

<210> 20
 <211> 7120
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> PI 痘mido Sint 騁 ico pXY152-endorf24-camtsr

<400> 20	
gaattcgtaa tcatgtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca	60
cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaata agtgagctaa	120
ctcacattaa ttgcgttgcg ctcaactgccc gctttccagt cgggaaacct gtctgtccag	180
ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtatttg gcgctcttcc	240
gcttctctgc tcaactgactc gctgcgctcg gtctttcggc tgcggcgagc ggtatcagct	300
cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg	360
tgagcaaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc	420
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga	480
aaccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct	540
cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg	600
gcgctttctc atagctcacg ctgtaggtat ctcaagtcgg ttaggtcgt tcgctccaag	660
ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttata cggttaactat	720
cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac	780

aggattagca gagcgaggta ttaggcggt gctacagagt tcttgaagtg gtggcctaac	840
tacggctaca ctagaagaac agtatttgggt atctgcgctc tgctgaagcc agttaccttc	900
ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggtttt	960
tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc	1020
ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg	1080
agattatcaa aaaggatcctt cacctagatc cttttggttc atgtgcagct ccatcagcaa	1140
aaggggatga taagtttatac accaccgact atttgcaaca gtgccgttga tcgtgctatg	1200
atcgactgat gtcacgacgc gtggagtgc atgtcgtgca atacgaatgg cgaaaagccg	1260
agctcaataa ctgccttaaa aaaattacgc cccgccctgc cactcatcgc agtactgttg	1320
taattcatta agcattctgc cgacatggaa gccatcacia acggcatgat gaacctgaat	1380
cgccagcggc atcagcacct tgcgccttg cgtataatat ttgcccatcg tgaaaacggg	1440
ggcgaagaag ttgtccatat tggccacgtt taagtcaaaa ctggtgaaac tcaccaggg	1500
attggctgag acgaaaaaca tattctcaat aaacccttta gggaaatagg ccaggttttc	1560
accgtaacac gccacatcctt gcgaatatat gtgtagaaac tgccggaaat cgtcgtggta	1620
ttcactccag agcgatgaaa acgtttcagt ttgctcatgg aaaacgggtg aacaagggtg	1680
aacactatcc catatcacca gtcaccgtc ttctattgcc atacgaaatt ccgatgagc	1740
attcatcagg cgggcaagaa tgtgaataaa ggccggataa aacttgtgct tatttttctt	1800
tacggtcttt aagaaggccg taatatccag ttgaacggtc tggttatagg tacattgagc	1860
aactgactga aatgcctcaa aatgttcttt acgatgccat tgggatatat caacggtggt	1920
atatccagtg atttttttct ccattttagc ttcccttagct cctgaaaatc tcgataactc	1980
aaaaaatacg cccggtagtg atcttatttc attatggtga aagttggaac ctcttacgtg	2040
ccgatcaacg tctcattttc gccaaaagtt ggcccagggc ttcccgttat caacagggac	2100
accaggattt atttattctg cgaagtgatc ttccgtcaca ggtattttatt cgcgataagc	2160

tcattggagcgc gcgtaaccgt cgcacaggaa ggacagagaa agcgcggaag cttattttaa	2220
taccgcgcgg gtcccggacg gggaagagcg gggagctttg ccagagagcg acgacttccc	2280
cttgcgttgg tgattgccgg tcagggcagc catccgccat cgtcgcgtag ggtgtcacac	2340
cccaggaatc gcgtcactga acacagcagc cggtaggacg accatgactg agttggacac	2400
catcgcaaat ccgtccgac cggcgggtgca gcgatcatc gatgtcacca agccgtcgcg	2460
atccaacata aagacaacgt tgatcgagga cgtcgagccc ctcatgcaca gcatcgcggc	2520
cggggtggag ttcatcgagg tctacggcag cgacagcagt ccttttccat ctgagttgct	2580
ggatctgtgc gggcggcaga acataccggt ccgcctcatc gactcctcga tcgtcaacca	2640
gttgttcaag ggggagcgga aggccaaagac attcggcatc gcccgctcc ctgcccggc	2700
caggttcggc gacatcgca gccggcgtgg ggacgtcgtc gttctcgac gggatgaagat	2760
cgtcgggaac atcggcgcga tagtacgcac gtcgctcgcg ctccggagcgt cggggatcat	2820
cctggtcgac agtgacatca ccagcatcgc ggaccggcgt ctccaaaggg ccagccgagg	2880
ttacgtcttc tcccttccc tcgttctctc cggtcgcgag gaggccatcg ctttcattcg	2940
ggacagcggg atgcaactga tgacgtcaa ggccgatggc gacatttccg tgaaggaaact	3000
cggggacaat ccggatcggc tggccttgct gttcggcagc gaaaagggtg ggoccttccga	3060
cctgttcgag gaggcgtctt ccgcctcggg ttccatcccc atgatgagcc agaccgagtc	3120
tctcaacgtt tccgtttccc tcggaatcgc gctgcacgag aggatcgaca ggaatctcgc	3180
ggccaaccga taagctagct agagtcgacc tgcaggtccc cggggatcgg tcttgccttg	3240
ctcgtcggtg atgtacttca ccagctccgc gaagtcgctc ttcttgatgg agcgcattggg	3300
gacgtgcttg gcaatcacgc gcacccccg gccgttttag cggctaataaa agtcatggct	3360
ctgcctcgg gcggaccacg cccatcatga ccttgccaag ctcgtcctgc ttctcttcga	3420
tcttcgccag cagggcgagg atcgtggcat caccgaaccg gccgtgcgc gggctcgtcg	3480
tgagccagag tttcagcagg ccgccaggc ggcccaggtc gccattgatg cgggccagct	3540
cgcggacgtg ctcatagtcc acgacgccg tgattttgta gccctggccg acggccagca	3600

ggtaggccga caggctcatg ccggccgccg ccgccttttc ctcaatcgct ctlogttcgt	3660
ctggaaggca gtacaccttg ataggtgggc tgcccttcct ggttggcttg gtttcatcag	3720
ccatccgctt gccctcatct gttacgccgg cggtagccgg ccagcctcgc agagcaggat	3780
tcccgttgag caccgccagg tgccaataag ggacagtga gaaggaacac ccgctcgcgg	3840
gtgggcctac ttacctatc ctgccggct gacgccgttg gatacaccaa ggaaagtcta	3900
cacgaaccct ttggcaaaat cctgtatatc gtgcgaaaaa ggatggatat accgaaaaaa	3960
togctataat gacccgaag cagggttatg cagcggaaaa gatccgtcga cctgcaggca	4020
tgcaagctct agcgattcca gacgtccga aggcgtggcg cggcttcccc gtgccggagc	4080
aatcgccctg ggtgggttac acgacgccc tctatggccc gtactgacgg acacaccgaa	4140
gccccggcgg caaccctcag cggatgcccc ggggcttcac gttttcccag gtcagaagcg	4200
gttttcggga gtagtgcccc aactggggta acctttgagt tctctcagtt gggggcgtag	4260
ggtcgccgac atgacacaag gggttgtgac cggggtggac acgtacgcgg gtgcttacga	4320
ccgtcagtcg cgcgagcgcg agaactcgag cgcagcaagc ccagcgacac agcgtagcgc	4380
caacgaagac aaggcggccg accttcagcg cgaagtcgag cgcgacgggg gccggttcag	4440
gttcgtcggg catttcagcg aagcgccggg cacgtcggcg ttcgggacgg cggagcgccc	4500
ggagttcgaa cgcattcctga acgaatgccg cgcggggcgg ctcaacatga tcattgtcta	4560
tgacgtgtcg cgcttctcgc gcctgaaggt catggacgcg attccgattg tctcggaatt	4620
gctcgccctg ggctgacga ttgtttccac tcaggaaggc gtcttcggcg agggaaacgt	4680
catggacctg attcacctga ttatgcggct cgacgcgtcg cacaagaat cttcgctgaa	4740
gtcggcgaag atttctgaca cgaagaacct tcagcgcgaa ttgggogggt acgtcggcgg	4800
gaaggcgctt tacggcttcg agcttgtttc ggagacgaag gagatcacgc gcaacggccg	4860
aatggtcaat gtcgtcatca acaagcttgc gcactcgacc actcccccta ccggaccctt	4920
cgagttcgag cccgacgtaa tccggtggtg gtggcgtgag atcaagacgc acaaacacct	4980

tcccttcaag ccgggcagtc aagccgccat tcacccgggc agcatcacgg ggctttgtaa	5040
gogcatggac gctgacgcgc tgccgacccg gggcgagacg attgggaaga agaccgcttc	5100
aagcgcttgg gaccgggcaa ccgttatgcg aatccttcgg gaccgcgta ttgcgggctt	5160
cgccgctgag gtgatctaca agaagaagcc ggacggcacg ccgaccacga agattgaggg	5220
ttaccgcatt cagcgcgacc cgatcacgct ccggccggtc gagcttgatt gcggaccgat	5280
catcgagccc gctgagtggc atgagcttca ggctggttg gacggcaggg ggcgcgcaa	5340
ggggctttcc cgggggcaag ccattctgtc cgccatggac aagctgtact gcgagtgtgg	5400
cgccgtcatg acttcaagc gcggggaaga atcgatcaag gactcttacc gctgccgtcg	5460
ccggaaggct gtcgaccgt ccgcacctgg gcagcacgaa ggcacgtgca acgtcagcat	5520
ggcggcactc gacaagttcg ttgcggaac catcttcaac aagatcaggc acgccgaagg	5580
cgacgaagag acgttggcgc ttctgtggga agccgcccga cgcttcggca agctcactga	5640
ggcgccctgag aagagcggcg aacgggcgaa ccttgttgcg gagcgcgcg acgccctgaa	5700
cgcccttgaa gagctgtacg aagaccgcgc ggcaggcgcg tacgacggac ccgttggcag	5760
gaagcacttc cggaagcaac aggcagcgct gacgctccgg cagcaagggg cggaagagcg	5820
gcttgccgaa cttgaagccg ccgaagcccc gaagcttccc cttgaccaat ggttccccga	5880
agacgccgac gctgaccga ccggccctaa gtcgtggtgg gggcgcgct cagtagacga	5940
caagcgcgtg ttctcgggc tcttcgtaga caagatcgtt gtcacgaagt cgactacggg	6000
cagggggcag ggaacgccc tcgagaagcg cgcttcgac acgtgggcga agccgccgac	6060
cgacgacgac gaagacgacg ccagggacgg cacggaagac gtagcggcgt aggtgagtga	6120
agatctagac gtggcaccgc gatgctgttg tgggcacaat cgtgccggtt ggtaggatcc	6180
acatatggaa ataagttcgc tctccaccga cggctcccc cggtacgac gggagagtcc	6240
cgagcacgtg gaaatgctgg ccgccgccga caccgcgtt ccaccgatca tggcgcaccg	6300
ccgcaccggg cgggtcatcg acggcatgca ccggctgcgc gccgcgatgc tgacgggccg	6360
tacgacgac gcggtgaggt tcttcgacgg caccgaggag gacgccttcg tcctcgccgt	6420

gaagtcgaac atcgcgacg gactgccgct gtccgccgcc gaccgccggc gggccgccgg	6480
gcgcatcatg gccaccatc cccggtggtc ggaccggatg atcgccctcg tggtcggcac	6540
ctccgccagg acggtcgccg agatccgccg cgacgccggc gccgccgggg cgggggagcc	6600
caccgcacg gccccgggacg gcagggtacg gcccgtcgac gtgagcgagg gccgcagact	6660
ggcccacgac atgatcgctc gcgaccggg cctgtcgtg cgccaggtcg ccgcgccgc	6720
cgggatctcg ccggagaccg tcagggacgt cagacaccgg atgctccgcg gtgaggacc	6780
ggtgcccgcg ccgcggccgc ggaccctggt ggagcgcggc gcggaccgcc gggcggagcc	6840
ggccgggaag gccgccgcgc cgtcggggac ggagccgccg ccgccgtcg tgatgaagcg	6900
gctgagggcc gatccggcg tgcgtctcaa cgagaacgga cgcgacctgc tgcggcttct	6960
ggatatccac acggtccggc tggaggactg gaaccgcatt atcgaaagcg tgccgccgca	7020
ccgtctggag acggtggcg agctggcacg ctctgcgcc gacaaatggt ccgagatcg	7080
gtcacgcac gaaagcaac catcacatct ggccgggtga	7120

<210> 21
 <211> 30
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 21	
gtggcaattt aaatggaaat gtgcgcggaa	30

<210> 22
 <211> 31
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 22

tatatagagc tcaacttggt ctgacagtta c

31

<210> 23

<211> 7162

<212> ADN

<213> Secuencia Artificial

<220>

<223> pXY152-endorf24-blatsr sint 驕 ico

<400> 23

gaattcgtaa tcatgtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca	60
cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaata agtgagctaa	120
ctcacattaa ttgcgttgcg ctcaactgccc gctttccagt cgggaaacct gtctgtccag	180
ctgcattaat gaatcggcca acgcgcgggg agaggcgggt tgcgtattgg gcgctcttcc	240
gcttcctcgc tcaactgactc gctgcgctcg gtctgttcggc tgcggcgagc ggtatcagct	300
cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg	360
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggogtttttc	420
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga	480
aaccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct	540
cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg	600
gcgctttctc atagctcacg ctgtaggtat ctcagttcgg ttaggtcgt tcgctccaag	660
ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttata cggtaactat	720
cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac	780
aggattagca gagcgaggta ttaggcgggt gctacagagt tcttgaagtg gtggcctaac	840
tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc agttacctc	900
ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggttt	960
tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc	1020
ttttctacgg ggtctgacgc tcagtgaac gaaaactcac gttaagggat tttggtcatg	1080

agattatcaa aaaggatcctt cacctagatc cttttggttc atgtgcagct ccatcagcaa	1140
aaggggatga taagtttatc accaccgact atttgaaca gtgccgttga tcgtgctatg	1200
atcgactgat gtcacacgag gtggagtga atgtcgtgca atacgaatgg cgaaaagccg	1260
agctcaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc	1320
tgtctatttc gttcatccat agttgcctga ctccccgtcg ttagataaac tacgatacgg	1380
gagggttac catctggccc cagtgcctga atgataccgc gagaccacg ctcaccggct	1440
ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggctcctgca	1500
actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg	1560
ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgttgt gtcacgctcg	1620
tcgtttggtg tggcttcatt cagctccggt tccaacgat caaggcgagt tacatgatcc	1680
cccatgttgt gcaaaaaagc ggtagctcc ttccgtcctc cgatcgttgt cagaagtaag	1740
ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg	1800
ccatccgtaa gatgcttttc tgtgacttgt gactactcaa ccaagtcatt ctgagaatag	1860
tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac cgcgccacat	1920
agcagaactt taaaagtgtc catcattgga aaacgttctt cggggcgaaa actctcaagg	1980
atcttaccgc tgttgagatc cagttcgatg taaccactc gtgcacccaa ctgatcttca	2040
gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca	2100
aaaaagggaa taaggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat	2160
tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag	2220
aaaaataaac aaataggggt tccgcgcaca ttccattta aataccgcgc gggccccgga	2280
cggggaagag cggggagctt tgccagagag cgacgacttc cccttgcgtt ggtgattgcc	2340
ggtcagggca gccatccgcc atcgtcgcgt aggggtgtcac accccaggaa tcgcgtcact	2400
gaacacagca gccggtagga cgaccatgac tgagttggac accatcgcaa atccgtccga	2460

tccggcggtg cagcggatca tcgatgtcac caagccgtcg cgatccaaca taaagacaac	2520
gttgatcgag gacgtcgagc ccctcatgca cagcatcgcg gccgggggtgg agttcatcga	2580
ggtctacggc agcgacagca gtccctttcc atctgagttg ctggatctgt gcgggcggca	2640
gaacataccg gtccgcctca tcgactcctc gatcgtcaac cagttgttca agggggagcg	2700
gaaggccaag acattcgga tcgcccgcgt ccctcgcccg gccaggttcg gcgacatcgc	2760
gagccggcgt ggggacgtcg tcgttctcga cggggtgaag atcgtcggga acatcggcgc	2820
gatagtacgc acgtcgctcg cgctcggagc gtcggggatc atcctggtcg acagtacat	2880
caccagcatc gcggaccggc gtctccaaag ggccagccga ggttacgtct tctcccttcc	2940
cgtcgttctc tccggtcgcg aggaggccat cgccttcatt cgggacagcg gtatgcaact	3000
gatgacgctc aaggcggatg gcgacatttc cgtgaaggaa ctcggggaca atccggatcg	3060
gctggccttg ctgttcggca gcgaaaaggg tgggccttcc gacctgttcg aggaggcgtc	3120
ttccgcctcg gtttccatcc ccatgatgag ccagaccgag tctctcaacg tttccgttcc	3180
cctcggaaac gcgctgcacg agaggatcga caggaatctc gcggccaacc gataagctag	3240
ctagagtcca cctgcaggtc cccggggatc ggtcttgcct tgctcgtcgg tgatgtactt	3300
caccagctcc gcgaagtgcg tcttcttgat ggagcgcatg gggacgtgct tggcaatcac	3360
gcgcaccccc cggccgtttt agcggctaaa aaagtcatgg ctctgccttc gggcggacca	3420
cgcccatcat gacctgccca agctcgtcct gcttctcttc gatcttcgcc agcagggcga	3480
ggatcgtggc atcacccaac cgcgccgtgc gcgggtcgtc ggtgagccag agtttcagca	3540
ggccgcccag gcggcccagg tcgccattga tgcgggccag ctgcgggacg tgctcatagt	3600
ccacgacgcc cgtgattttg tagccctggc cgacggccag caggtaggcc gacaggctca	3660
tgccggccgc cgccgccttt tcctcaatcg ctcttcgttc gtctggaagg cagtacacct	3720
tgataggtgg gctgcccttc ctggttggct tggtttcata agccatccgc ttgccctcat	3780
ctgttacgcc ggcggtagcc ggccagcctc gcagagcagg attccggtt agcaccgcca	3840
ggtgcgaata agggacagtg aagaaggaac acccgctcgc ggtgggcct acttcaccta	3900

tccgtcccgg ctgacgccgt tggatacacc aaggaaagtc tacacgaacc ctttggcaaa	3960
atcctgtata tcgtgcgaaa aaggatggat ataccgaaaa aatcgctata atgaccccgga	4020
agcagggtta tgcagcggaa aagatccgtc gacctgcagg catgcaagct ctagcgattc	4080
cagacgtccc gaagcgtgg cgcggcttcc ccgtgccgga gcaatcgccc tgggtgggtt	4140
acacgacgcc cctctatggc ccgtactgac ggacacaccg aagccccggc ggcaaccctc	4200
agcggatgcc ccggggcttc acgttttccc aggtcagaag cggttttcgg gagtagtgcc	4260
ccaactgggg taacctttga gttctctcag ttgggggcgt aggttcgccc acatgacaca	4320
aggggttttg accgggggtg acacgtacgc ggggtgcttac gaccgtcagt cgcgcgagcg	4380
cgagaactcg agcgcagcaa gcccgcgac acagcgtagc gccaacgaag acaaggcggc	4440
cgaccttcag cgcgaagtcg agcgcgacgg gggccggttc aggttcgtcg ggcatttcag	4500
cgaagcgccg ggcacgtcgg cgttcgggac ggcggagcgc ccggagttcg aacgcaccc	4560
gaacgaatgc cgcgcggggc ggctcaacat gatcattgtc tatgacgtgt cgcgcttctc	4620
gcgcctgaag gtcattggacg cgattccgat tgtctcgaa ttgctcgccc tgggcgtgac	4680
gattgtttcc actcaggaag gcgtcttcg gcagggaaac gtcattggacc tgattcacct	4740
gattatgcgg ctgcacgcgt cgcacaaaga atcttcgctg aagtcggcga agattctcga	4800
cacgaagaac cttcagcgcg aattgggcgg gtacgtcggc gggaaggcgc cttacggctt	4860
cgagcttggt tcggagacga aggagatcac gcgcaacggc cgaatggtca atgtcgtcat	4920
caacaagctt gcgcactcga ccaactcccct taccggaccc ttcgagttcg agcccgacgt	4980
aatccggtgg tgggtggcgtg agatcaagac gcacaaacac cttcccttca agccgggcag	5040
tcaagccgcc attcaccogg gcagcatcac ggggctttgt aagcgcattg acgctgacgc	5100
cgtgccgacc cggggcgaga cgattgggaa gaagaccgt tcaagcgctt gggacccggc	5160
aaccgttatg cgaatccttc gggacccgcg tattgcgggc ttcccgctg aggtgatcta	5220
caagaagaag ccggacggca cgcggaccac gaagattgag ggttacgca ttcagcgca	5280

cccgatcacg ctccggccgg tcgagcttga ttgcggaccg atcatcgagc ccgctgagtg	5340
gtatgagctt caggcgttgt tggacggcag ggggcgcggc aaggggcttt ccggggggca	5400
agccattctg tccgccatgg acaagctgta ctgcgagtg ggcccgctca tgacttcgaa	5460
gcgcggggaa gaatcgatca aggactctta ccgctgccgt cgccggaagg tggtcgaccc	5520
gtccgcacct gggcagcacg aaggcacgtg caacgtcagc atggcggcac tcgacaagtt	5580
cgttgcggaa cgcatcttca acaagatcag gcacgccgaa ggcgacgaag agacgttggc	5640
gcttctgttg gaagccgccc gacgcttcgg caagctcact gaggcgctg agaagagcgg	5700
cgaacgggcg aaccttggtg cggagcgcgc cgacgccctg aacgcccttg aagagctgta	5760
cgaagaccgc gcggcaggcg cgtacgacgg acccgttggc aggaagcact tccggaagca	5820
acaggcagcg ctgacgtccc ggcagcaagg ggcggaagag cggcttgccg aacttgaagc	5880
cgccgaagcc ccgaagcttc cccttgacca atggttcccc gaagacgccg acgctgaccc	5940
gaccggccct aagtcgttgt gggggcgcgc gtcagtagac gacaagcgcg tgttcgtcgg	6000
gctcttcgta gacaagatcg ttgtcacgaa gtcgactacg ggcagggggc agggaaacgcc	6060
catcgagaag cgcgcttcga tcacgtgggc gaagccgccg accgacgacg acgaagacga	6120
cgcccaggac ggcacggaag acgtagcggc gtaggtgagt gaagatctag acgtggcacc	6180
gcgatgctgt tgtgggcaca atcgtgccgg ttggtaggat ccacatatgg aaataagttc	6240
gctctccacc gacggctccc cgcggatcga cggggagagt cccgagcacg tggaaatgct	6300
ggccgccgcc gacaccgcgc ttccaccgat catggtgcac cgccgcaccg ggcgggtcat	6360
cgacggcatg caccggctgc gcgccgcgat gctgacgggc cgtacgacga tcgcggtgag	6420
gttcttcgac ggcaccgagg aggacgcctt cgtcctcgcc gtgaagtcga acatcgcgca	6480
cggactgccg ctgtccgccg ccgaccgccg gcgggccgcc gggcgcatca tggccaccca	6540
tccccggttg tcggaccgga tgatgcctc ggtggtcggc acctccgccg ggacggtcgc	6600
cgagatccgc cgcgacgccg gcgccgccgg ggcgggggag cccacccgca tcggccggga	6660
cggcagggta cggcccgtcg acgtgagcga gggccgcaga ctggcccacg acatgatcgt	6720

cgcgcacccg ggctgtgc tgcgcaggt cgcgcgcgc gccgggatct cgcgcgagac	6780
cgtcaggac gtcagacacc ggatgctccg cggtaggac cgggtgccg cgcgcggcc	6840
gcgaccctg gtggagcgc gcgcggaccg ccgggcggag ccggccggga aggcgcgcgc	6900
gccgtgcggg acggagccgc cgcgcgcgt cgtgatgaag cggctgagg ccatccgc	6960
gctgcgtctc aacgagaacg gacgcacct gctcggctt ctggatatcc acacgtccg	7020
gctggaggac tggaaccgca ttatcgaaag cgtgccgcg caccgtctgg agacgtggc	7080
gcagctggca cgctcctgc cgcacaaatg gtccgagatc gcgtcacgca tcgaaagca	7140
cgcacacat ctggccgggt ga	7162

<210> 24
 <211> 24
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Gebador de Oligonucle · ido Sint 騁 ico

<400> 24	
aaggagaaga gccttcagaa ggaa	24

<210> 25
 <211> 321
 <212> PRT
 <213> S. griseus

<400> 25

Met	Asp	Pro	Thr	Arg	Val	Asp	Ile	Phe	Ala	Leu	Pro	Ala	Val	Glu	Ile
1				5					10					15	

Glu	Leu	Ser	Arg	Leu	Ser	Ser	Ala	Ser	Ser	Pro	Arg	Thr	Ser	Gly	Glu
			20					25					30		

Asp	Pro	Glu	His	Val	Glu	Thr	Leu	Leu	Ser	Ala	Glu	Gly	Glu	Leu	Pro
		35					40					45			

Pro Ile Leu Val His Arg Pro Thr Met Gln Val Leu Asp Gly Leu His
50 55 60

Arg Leu Lys Val Ala Arg Val Arg Gly Asp Thr Lys Ile Leu Ala Arg
65 70 75 80

Leu Val Asp Ala Thr Glu Ser Asp Ala Phe Val Leu Ala Val Glu Ala
85 90 95

Asn Ile Arg His Gly Leu Pro Leu Ser Leu Ala Asp Arg Lys Arg Ala
100 105 110

Ala Val Gln Ile Ile Gly Thr His Pro Gln Trp Ser Asp Arg Arg Val
115 120 125

Ala Ser Ala Thr Gly Ile Ser Ala Gly Thr Val Ala Asp Leu Arg Arg
130 135 140

Arg Ala Gly Glu Asp Gly Thr Glu Ala Arg Ile Gly Arg Asp Gly Arg
145 150 155 160

Val Arg Pro Ser Asp Gly Ser Glu Arg Arg Arg Leu Ala Ala Glu Leu
165 170 175

Ile Arg Ser Asp Pro Gly Leu Ser Leu Arg Gln Val Ala Lys Gln Val
180 185 190

Gly Ile Ser Pro Glu Thr Val Arg Asp Val Arg Gly Arg Leu Glu Arg
195 200 205

Gly Glu Ser Pro Thr Pro Asp Gly Thr Arg Arg Leu Pro Ala Lys Pro
210 215 220

His Pro Leu Arg Leu Ser Glu Pro Asp Phe Gly Arg Ala Val Asp Gln

225 230 235 240

Asp Arg Leu Ala Leu Leu Glu Arg Leu Lys Ser Asp Pro Ala Leu Arg
 245 250 255

Leu Asn Glu Val Gly Arg Ile Leu Leu Arg Met Leu Thr Met His Ser
 260 265 270

Met Asp Gly Gln Glu Trp Glu Arg Ile Leu Gln Gly Val Pro Pro His
 275 280 285

Leu His Gly Val Ile Ala Gly Phe Ala Arg Asp His Ala Arg Val Trp
 290 295 300

Ala Glu Phe Ala Asp His Leu Glu Ser Arg Ala Thr Glu Leu Ala Ala
305 310 315 320

Gly

<210> 26
<211> 311
<212> PRT
<213> Streptomyces fungicidicus

<400> 26

Val Glu Ile Ser Ser Leu Ser Thr Asp Gly Ser Pro Arg Ile Asp Gly
1 5 10 15

Glu Ser Pro Glu His Val Glu Met Leu Ala Ala Ala Asp Thr Ala Leu
20 25 30

Pro Pro Ile Met Val His Arg Arg Thr Gly Arg Val Ile Asp Gly Met
35 40 45

His Arg Leu Arg Ala Ala Met Leu Thr Gly Arg Thr Thr Ile Ala Val

50	55	60
Arg Phe Phe Asp Gly Thr Glu Glu Asp Ala Phe Val Leu Ala Val Lys		
65	70	75 80
Ser Asn Ile Ala His Gly Leu Pro Leu Ser Ala Ala Asp Arg Arg Arg		
	85	90 95
Ala Ala Gly Arg Ile Met Ala Thr His Pro Arg Trp Ser Asp Arg Met		
	100	105 110
Ile Ala Ser Val Val Gly Thr Ser Ala Arg Thr Val Ala Glu Ile Arg		
	115	120 125
Arg Asp Ala Gly Ala Ala Gly Ala Gly Glu Pro Thr Arg Ile Gly Arg		
	130	135 140
Asp Gly Arg Val Arg Pro Val Asp Val Ser Glu Gly Arg Arg Leu Ala		
	145	150 155 160
His Asp Met Ile Val Arg Asp Pro Gly Leu Ser Leu Arg Gln Val Ala		
	165	170 175
Arg Ala Ala Gly Ile Ser Pro Glu Thr Val Arg Asp Val Arg His Arg		
	180	185 190
Met Leu Arg Gly Glu Asp Pro Val Pro Ala Pro Arg Pro Arg Thr Leu		
	195	200 205
Val Glu Arg Gly Ala Asp Arg Arg Ala Glu Pro Ala Gly Lys Ala Ala		
	210	215 220
Ala Pro Cys Gly Thr Glu Pro Pro Pro Ala Val Val Met Lys Arg Leu		
	225	230 235 240

Arg Ala Asp Pro Ala Leu Arg Leu Asn Glu Asn Gly Arg Asp Leu Leu
245 250 255

Arg Leu Leu Asp Ile His Thr Val Arg Leu Glu Asp Trp Asn Arg Ile
260 265 270

Ile Glu Ser Val Pro Pro His Arg Leu Glu Thr Val Ala Gln Leu Ala
275 280 285

Arg Ser Cys Ala Asp Lys Trp Ser Glu Ile Ala Ser Arg Ile Glu Ser
290 295 300

Asn Ala Ser His Leu Ala Gly
305 310

<210> 27
<211> 15
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de Oligonucle · ido Sint 騁 ico

<400> 27
gtgtttaatt aatga

15

<210> 28
<211> 321
<212> PRT
<213> Amycolatopsis

<400> 28

Met Asp Pro Thr Arg Val Asp Ile Phe Ala Leu Pro Ala Val Glu Ile
1 5 10 15

Glu Leu Ser Arg Leu Ser Ser Ala Ser Ser Pro Arg Thr Ser Gly Glu
20 25 30

Asp Pro Glu His Val Glu Thr Leu Leu Ser Ala Glu Gly Glu Leu Pro
35 40 45

Pro Ile Leu Val His Arg Pro Thr Met Gln Val Leu Asp Gly Leu His
50 55 60

Arg Leu Lys Val Ala Arg Val Arg Gly Asp Thr Lys Ile Leu Ala Arg
65 70 75 80

Leu Val Asp Ala Thr Glu Ser Asp Ala Phe Val Leu Ala Val Glu Ala
85 90 95

Asn Ile Arg His Gly Leu Pro Leu Ser Leu Ala Asp Arg Lys Arg Ala
100 105 110

Ala Val Gln Ile Ile Gly Thr His Pro Gln Trp Ser Asp Arg Arg Val
115 120 125

Ala Ser Ala Thr Gly Ile Ser Ala Gly Thr Val Ala Asp Leu Arg Arg
130 135 140

Arg Ala Gly Glu Asp Gly Thr Glu Ala Arg Ile Gly Arg Asp Gly Arg
145 150 155 160

Val Arg Pro Ser Asp Gly Ser Glu Arg Arg Arg Leu Ala Ala Glu Leu
165 170 175

Ile Arg Ser Asp Pro Gly Leu Ser Leu Arg Gln Val Ala Lys Gln Val
180 185 190

Gly Ile Ser Pro Glu Thr Val Arg Asp Val Arg Gly Arg Leu Glu Arg
195 200 205

Gly Glu Ser Pro Thr Pro Asp Gly Thr Arg Arg Leu Pro Ala Lys Pro
210 215 220

His Pro Leu Arg Leu Ser Glu Pro Asp Phe Gly Arg Ala Val Asp Gln
225 230 235 240

Asp Arg Leu Ala Leu Leu Glu Arg Leu Lys Ser Asp Pro Ala Leu Arg
245 250 255

Leu Asn Glu Val Gly Arg Ile Leu Leu Arg Met Leu Thr Met His Ser
260 265 270

Met Asp Gly Gln Glu Trp Glu Arg Ile Leu Gln Gly Val Pro Pro His
275 280 285

Leu His Gly Val Ile Ala Gly Phe Ala Arg Asp His Ala Arg Val Trp
290 295 300

Ala Glu Phe Ala Asp His Leu Glu Ser Arg Ala Thr Glu Leu Ala Ala
305 310 315 320

Gly

<210> 29
<211> 395
<212> PRT
<213> S. kasugaensis

<400> 29

Met Ala Glu Thr Val Arg Ala Asp Ser Pro Leu Lys Ser Ser Tyr Arg
1 5 10 15

Asn Val Pro Ala Ala Glu Val Gln Gly Ser Gly Leu Ser Val Gly Gln
20 25 30

Arg Thr Thr Arg Ile Ala Ile Ser Ser Leu Leu Ala Ala Asp Ser Pro
35 40 45

Arg Ser Ala Gly Glu Asn Ala Glu His Ile Arg Leu Leu Ala Asp Ser
50 55 60

Gly Ala Arg Leu Pro Pro Ile Val Val Gln Arg Ser Thr Met Arg Val
65 70 75 80

Ile Asp Gly Met His Arg Leu Arg Ala Ala Ala Leu Arg Gly Glu Thr
85 90 95

Glu Ile Glu Val Arg Phe Phe Asp Gly Ala Glu Glu Asp Ser Phe Leu
100 105 110

Leu Ala Val Arg Ser Asn Ile Ala His Gly Leu Pro Leu Ser Gln Glu
115 120 125

Glu Arg Ala Ala Ala Ala Gln Arg Ile Ile Arg Ser His Ala Gln Trp
130 135 140

Ser Asn Gln Ala Ile Gly Glu Val Thr Gly Leu Asp Ala Lys Thr Ile
145 150 155 160

Ala Ala Leu Arg Arg Asp Ala Lys Asp Val Pro Gln Leu Asp Ala Arg
165 170 175

Ile Gly Arg Asp Gly Arg Val Arg Pro Val Asp Gly Ala Gln Gly Arg
180 185 190

Arg Leu Ala Gly Glu Leu Met Ala Glu Gln Pro Asp Ala Pro Leu Arg
195 200 205

Lys Ile Ala His Ala Ala Gly Val Ser Leu Gly Thr Ala Ser Asp Val
210 215 220

Arg Arg Arg Ile Arg Asn Gly Gln Asp Pro Val Pro Ala Gly Arg Gln
225 230 235 240

Lys Ala Asp Pro Gln Pro Pro Ala Arg Tyr Ala Ala Ser Glu Asp Arg
245 250 255

Ser Gly Thr Thr Ala Pro Arg Thr Gly Glu Gln Asn Arg Arg Val Leu
260 265 270

Leu Gln Lys Leu Arg Lys Asp Pro Ser Leu Arg Cys Asn Glu Ala Gly
275 280 285

Arg Ala Leu Leu Arg Trp Leu Glu Val Gln Ala Val Glu Gly Glu Asp
290 295 300

Trp Glu Arg Leu Leu Asp Ser Val Pro Met His Cys Ala Ala Thr Ile
305 310 315 320

Val Glu Leu Ala Arg Arg Lys Asp Pro Ser Leu Arg Cys Asn Glu Ala
325 330 335

Gly Arg Ala Leu Leu Arg Trp Leu Glu Val Gln Ala Val Glu Gly Glu
340 345 350

Asp Trp Glu Arg Leu Leu Asp Ser Val Pro Met His Cys Ala Ala Thr
355 360 365

Ile Val Glu Leu Ala Arg Gly Cys Ser Gly Val Trp Gln Asp Phe Ala
370 375 380

Ala Gln Leu Glu Arg Arg Gly Arg Ala Ser Ala
385 390 395

<210> 30
<211> 367
<212> PRT
<213> S. niveus

<400> 30

Met Thr Asn Ser Gly Asp Glu Glu Ile Thr Pro Ala Ser Leu Lys Ala
1 5 10 15

Thr Arg Lys Gly Glu Arg Val Ser Ile Gly Ser Leu Leu Pro Pro Ser
20 25 30

Glu Leu Val Arg Ser Gly Glu Ser Thr Glu His Ile Arg Val Leu Ala
35 40 45

Glu Thr Asp Glu Asp Leu Pro Pro Ile Val Val His Arg Gly Thr Arg
50 55 60

Arg Val Val Asp Gly Met His Arg Leu Trp Ala Ala Arg Phe Arg Gly
65 70 75 80

Asp Glu Ser Ile Glu Val Val Phe Val Asp Gly Ser Pro Ala Asp Val
85 90 95

Phe Val Leu Ala Val Glu Leu Asn Arg Ala His Gly Leu Pro Leu Thr
100 105 110

Leu Asp Glu Arg Lys Ser Ala Ala Ala Gln Ile Met Asp Ser His Pro
115 120 125

His Trp Ser Asp Arg Lys Ile Ala Arg Thr Thr Gly Leu Ala Ala Ser
130 135 140

Thr Val Ala Ser Leu Arg Ser Ser Ser Thr Ala Gly Thr Val Gly Arg
145 150 155 160

Arg Thr Gly Gln Asp Gly Arg Ser Arg Pro Asn Asp Gly Thr Asp Gly
165 170 175

Arg Gln Arg Ala Ala Ala Leu Leu Ala Arg Asn Pro Asn Ala Ser Leu

180	185	190
Arg Glu Val Thr Arg Ala Ala Gly Ile Ser Val Gly Thr Ala Ser Asp		
195	200	205
Val Arg Ala Arg Leu Arg Arg Gly Glu Pro Ala Leu Thr Ala Arg Gln		
210	215	220
Gln Ala Val Met Lys Leu Arg Pro Ala Ala Arg Asn Pro Asn Ala Ser		
225	230	235 240
Leu Arg Glu Val Thr Arg Ala Ala Gly Ile Ser Val Gly Thr Ala Ser		
245	250	255
Asp Val Arg Ala Arg Leu Arg Arg Gly Glu Pro Ala Leu Thr Ala Arg		
260	265	270
Gln Gln Ala Val Met Lys Leu Arg Pro Ala Ala Ala Gln Arg Ser Gly		
275	280	285
Pro Asp Tyr Gly Arg Val Leu Glu Asn Leu Arg Lys Asp Pro Ser Leu		
290	295	300
Arg Phe Thr Asp Leu Gly Arg Arg Leu Leu Arg Leu Leu Asp Gly Ser		
305	310	315 320
Val Pro Gly Ser Val Glu Gln Ile Ala Gln Ile Ala Asp Gly Val Pro		
325	330	335
Glu His Cys Arg Thr Val Val Val Asp Met Ala Arg Glu Cys Ala Ala		
340	345	350
Ala Trp Gln His Leu Ala Asp Gln Leu Ala Asp Arg Asp Thr Ala		
355	360	365

<210> 31
<211> 359
<212> PRT
<213> S. globisporus

<400> 31

Met Lys Ser Asp Ser Ala Gln Arg Ala Val Glu Arg Ser Arg Arg Val
1 5 10 15

Val Arg Ile Asp Glu Leu Ile Pro Ala Asp Ser Pro Arg Leu Asn Gly
20 25 30

Ile Asp Arg Ser His Val Gln Arg Leu Ala Thr Val Tyr Ala Ser Leu
35 40 45

Pro Pro Val Leu Val His Arg Pro Thr Met Arg Val Val Asp Gly Met
50 55 60

His Arg Ile Gly Ala Ala Arg Leu Lys Gly Leu Asp Thr Val Glu Val
65 70 75 80

Thr Phe Phe Glu Gly Ala Glu Glu Gln Val Phe Leu Arg Ser Val Ala
85 90 95

Ala Asn Ile Thr Asn Gly Leu Pro Leu Ser Val Ala Asp Arg Lys Thr
100 105 110

Ala Ala Ala Arg Ile Leu Ala Ser His Pro Thr Leu Ser Asp Arg Ala
115 120 125

Val Ala Ala His Val Gly Leu Asp Ala Lys Thr Val Ala Gly Val Arg
130 135 140

Thr Cys Ser Ala Ala Gly Ser Pro Leu Leu Asn Met Arg Thr Gly Ala
145 150 155 160

Asp Gly Arg Val His Pro Leu Asp Arg Thr Ala Glu Arg Leu His Ala
165 170 175

Ala Ala Leu Leu Thr Gln Asp Pro Gly Leu Pro Leu Arg Ser Val Val
180 185 190

Glu Gln Thr Gly Leu Ser Leu Gly Thr Ala His Asp Val Arg Arg Arg
195 200 205

Leu Leu Arg Gly Glu Asp Pro Val Pro Gln Asn Arg Gln Ser Ala Met
210 215 220

Leu Glu Pro Gly Leu Ala Pro Gln Lys Lys Ala Thr Ala Lys Pro Pro
225 230 235 240

Val Gly Pro Ala Ala Arg Pro Val Pro Lys Val Pro Pro Ala Val Ala
245 250 255

Gly Arg Pro Pro Val Ser Pro Arg Ser Arg Ala Pro Leu Glu Ala Leu
260 265 270

Arg Lys Leu Ser Asn Asp Pro Ser Leu Arg His Ser Asp Gln Gly Arg
275 280 285

Glu Leu Met Arg Trp Leu His Asn Arg Phe Val Val Asp Glu Ala Trp
290 295 300

Arg Arg Arg Ala Asp Ala Val Pro Ala His Cys Val Asp Ser Met Ala
305 310 315 320

Glu Leu Ala Gln His Cys Ser Asp Ala Trp His Arg Phe Ala Glu Glu
325 330 335

Met Val Arg Arg Arg His Ser Ala Ala Ala Asp Gly Ser Gly Leu Arg
340 345 350

Thr Thr Gln Pro Thr Arg Arg
355

<210> 32
<211> 329
<212> PRT
<213> Actinoplanes teichomyceticus

<400> 32

Met Thr Pro Asp Glu Glu Ala Leu Asn Arg Gln Pro Ile Met Glu Met
1 5 10 15

Glu Ile Ser Ser Leu Ser Leu Gly Gly Ser Pro Arg Leu Ala Gly Gly
20 25 30

Asp Pro Val His Leu Glu Ala Met Val Ala Ala Gln Gly Glu Leu Pro
35 40 45

Pro Ile Val Val His Arg Pro Thr Met Arg Val Ile Asp Gly Ser His
50 55 60

Arg Ile Gln Ala Ala Leu Arg Arg Gly Glu Thr Thr Ile Ala Gly Arg
65 70 75 80

Phe Phe Asp Gly Ser Asp Asp Glu Ala Phe Val Met Ser Val Trp Leu
85 90 95

Asn Val Ser His Gly Leu Pro Leu Ala Leu Ala Asp Arg Lys Arg Ala
100 105 110

Ala Glu Arg Ile Ala Val Ser His Pro Gln Trp Ser Asp Arg Arg Val
115 120 125

Ala Ala Val Thr Gly Ile Ser Pro Ser Thr Val Ala Asp Ile Arg Arg
130 135 140

Arg Val Ala Gly Thr Ser Ala Pro Glu Ala Ser Arg Ile Gly Gln Asp
145 150 155 160

Gly Arg Val Arg Pro Leu Asp Cys Ser Ala Gly Arg Leu Leu Ala Gly
165 170 175

Arg Leu Met Ala Glu Asn Pro Ala Leu Ser Leu Arg Gln Val Ala Lys
180 185 190

Ala Ala Ala Ile Ser Pro Glu Thr Ala Arg Asp Val Arg Asn Arg Leu
195 200 205

Leu Ser Gly Ala Glu Leu Val Pro Asn Arg Arg Pro Arg Asp Ala Ala
210 215 220

Pro Val Gly Val Lys Gly Gly Arg Asp Arg Arg Pro Leu Asn Leu Ile
225 230 235 240

Arg Ser Gly Asp Arg Pro Glu Pro Val Pro Asp His Ala Val Val Ile
245 250 255

Asn Arg Leu Met Ser Asp Pro Ala Leu Arg Tyr Thr Asp Thr Gly Arg
260 265 270

Asn Leu Leu Arg Leu Leu Ser Leu His Thr Arg Trp Ala Lys Glu Trp
275 280 285

Glu Ala Ile Val Asp Asn Leu Pro Pro His Cys Ala Asp Ala Val Ala
290 295 300

Asp Leu Ala Arg Gln Phe Ala Asp Leu Trp Ala Asp Phe Ala Ser Arg
305 310 315 320

Val Gly Pro Glu Glu Arg Met Ala Ser
325

<210> 33
<211> 221
<212> PRT
<213> S. coelicolor

<400> 33

Met Thr Ile Arg Leu Leu Ile Val Asp Asp Gln Glu Leu Ile Arg Thr
1 5 10 15

Gly Phe Arg Leu Phe Leu Gln Thr Gln Asn Asp Leu Glu Val Val Gly
20 25 30

Glu Ala Asp Asp Gly His Gly Ala Leu Ala Gln Ala Ala Ala Leu Arg
35 40 45

Pro Asp Val Val Leu Met Asp Ile Arg Met Pro Arg Met Asp Gly Val
50 55 60

Glu Ala Thr Ser Arg Leu Thr Ala Ser Asp Ser Pro Pro Arg Val Leu
65 70 75 80

Ile Leu Thr Thr Tyr Asp Leu Asp Glu Tyr Val Phe Gly Ala Leu Arg
85 90 95

Ala Gly Ala Ser Gly Phe Leu Leu Lys Asp Ala Ser Arg Asp Arg Leu
100 105 110

Leu Glu Ala Ile Arg Val Val His Ala Gly Glu Ala Leu Leu Ser Pro
115 120 125

Ser Ile Thr Arg Arg Leu Ile Glu Asp Tyr Ala Thr Arg Ala Ala Pro
130 135 140

Val Arg Pro Arg Glu Ala Val Leu Ala Gly Leu Thr Pro Arg Glu Arg
145 150 155 160

Glu Ile Leu Leu Leu Val Ala Arg Gly Leu Ser Asn Pro Glu Ile Ala
165 170 175

Ala Arg Leu Val Val Thr Glu Ala Thr Val Lys Ser His Val Gly Ser
180 185 190

Met Phe Ala Lys Leu His Leu Arg Asp Arg Ala Gln Ala Val Val Phe
195 200 205

Ala Tyr Glu Asn Ala Ile Val Leu Pro Gly Gly Thr Gly
210 215 220

<210> 34
<211> 222
<212> PRT
<213> S. coelicolor

<400> 34

Met Ile Arg Val Leu Leu Ala Asp Asp Glu Thr Ile Ile Arg Ala Gly
1 5 10 15

Val Arg Ser Ile Leu Thr Thr Glu Pro Gly Ile Glu Val Val Ala Glu
20 25 30

Ala Ser Asp Gly Arg Glu Ala Val Glu Leu Ala Arg Lys His Arg Pro
35 40 45

Asp Val Ala Leu Leu Asp Ile Arg Met Pro Glu Met Asp Gly Leu Thr
50 55 60

Ala Ala Gly Glu Met Arg Thr Thr Asn Pro Asp Thr Ala Val Val Val
65 70 75 80

Leu Thr Thr Phe Gly Glu Asp Arg Tyr Ile Glu Arg Ala Leu Asp Gln
85 90 95

Gly Val Ala Gly Phe Leu Leu Lys Ala Ser Asp Pro Arg Asp Leu Ile
100 105 110

Ser Gly Val Arg Ala Val Ala Ser Gly Gly Ser Cys Leu Ser Pro Leu
115 120 125

Val Ala Arg Arg Leu Met Thr Glu Leu Arg Arg Ala Pro Ser Pro Arg
130 135 140

Ser Glu Val Ser Gly Glu Arg Thr Thr Leu Leu Thr Lys Arg Glu Gln
145 150 155 160

Glu Val Leu Gly Met Leu Gly Ala Gly Leu Ser Asn Ala Glu Ile Ala
165 170 175

Gln Arg Leu His Leu Val Glu Gly Thr Ile Lys Thr Tyr Val Ser Ala
180 185 190

Ile Phe Thr Gln Leu Glu Val Arg Asn Arg Val Gln Ala Ala Ile Ile
195 200 205

Ala Tyr Glu Ala Gly Leu Val Lys Asp Ala Asp Leu Asn Arg
210 215 220

<210> 35
<211> 218
<212> PRT
<213> S. coelicolor

<400> 35

Met Arg Glu Asp Gly Lys Ile Arg Val Phe Leu Leu Asp Asp His Glu
1 5 10 15

Val Val Arg Arg Gly Val His Asp Leu Leu Ser Gly Glu Ala Asp Ile
20 25 30

Glu Val Val Gly Glu Ala Gly Thr Ala Ala Glu Ala Gln Ala Arg Val
35 40 45

Thr Ala Thr Arg Pro Asp Val Ala Val Leu Asp Val Arg Leu Pro Asp
50 55 60

Gly Ser Gly Val Glu Val Cys Arg Asp Ile Arg Ser Arg Asp Glu Ser
65 70 75 80

Val Arg Cys Leu Met Leu Thr Ser Phe Ala Asp Asp Glu Ala Leu Phe
85 90 95

Asp Ala Ile Met Ala Gly Ala Ser Gly Tyr Val Leu Lys Asp Ile Arg
100 105 110

Gly Ala Glu Leu Leu Gly Ala Val Arg Glu Val Ala Ala Gly Lys Ser
115 120 125

Leu Leu Asp Pro Ala Ala Thr Ala Arg Val Leu Glu Arg Leu Arg Gly
130 135 140

Gly Gly Ala Arg Pro Asp Asp Arg Leu Ala Arg Leu Thr Glu Gln Glu
145 150 155 160

Arg Arg Ile Leu Glu Leu Ile Gly Glu Gly Leu Thr Asn Arg Ala Ile
165 170 175

Gly Glu Arg Leu His Leu Ala Glu Lys Thr Ile Lys Asn Tyr Val Ser
180 185 190

Ser Leu Leu Gly Lys Leu Gly Met Gln Arg Arg Ser Gln Ala Ala Ala
195 200 205

Phe Val Ala Arg Leu Glu Ala Glu Asn Arg

210

215

<210> 36
<211> 220
<212> PRT
<213> Streptomyces fungicidicus

<400> 36

Val Ser Val Leu Leu Glu Gln Pro Ala Ser Leu Val Ala Tyr Arg Pro
1 5 10 15

Asn Lys Pro Thr Ala Met Val Val Val Ala Asp Pro Arg Val Arg Ser
20 25 30

Thr Val Thr Arg His Leu Trp Ala Leu Gly Val Arg Asp Val Ile Glu
35 40 45

Ala Ser Ser Val Ala Glu Ala Arg Pro Arg Ile Gly Asn Pro Arg Asp
50 55 60

Ile Cys Val Ala Glu Val His Leu Pro Asp Gly Ser Gly Leu Thr Leu
65 70 75 80

Leu Ser Glu Thr Arg Ala Ala Gly Trp Pro Asn Gly Leu Ala Leu Ser
85 90 95

Ala Ala Asp Asp Ile Gly Ala Val Arg Asn Ala Leu Ala Gly Gly Val
100 105 110

Lys Gly Tyr Val Val Thr Gly Thr Arg Thr Asn Leu Gly Leu Pro Thr
115 120 125

Arg Pro Gly Ala Ala Pro Ile Gly Ala Ala Ala Arg Leu His Arg
130 135 140

Arg Pro Pro Gly Ala Pro Ser His Pro Gly Gly Tyr Arg Glu Leu Ser

145 150 155 160

Lys Ala Ile Gly Val Ser Met Gly Leu Ser Ala Leu Thr Val Lys Ser
180 185 190

Met Val Ala Val Ala Leu Arg Thr Gly Ile Ile His
210 215 220

tga

663

<210> 38

<211> 936

<212> ADN

<213> Streptomyces fungicidicus

<400> 38

gtggaaataa gttcgtctc caccgacggc tccccgcgga tcgacgggga gagtcccgag	60
cacgtggaaa tgctggccgc cgccgacacc gcgcttccac cgatcatggt gcaccgccgc	120
accgggcggg tcatcgacgg catgcaccgg ctgcgcgcgc cgatcgtgac gggccgtacg	180
acgatcgagg tgaggttctt cgacggcacc gaggaggacg cttcgtcct cgccgtgaag	240
tcgaacatcg cgcacggact gccgtgtcc gccgccgacc gccggcgggc cgccgggcgc	300
atcatggcca cccatccccg gtggtcggac cggatgatcg cctcgggtgt cggcacctcc	360
gccaggacgg tcgccgagat ccgccgcgac gccggcgcgc ccggggcggg ggagcccacc	420
cgcacggcc gggacggcag ggtacggccc gtcgacgtga gcgagggcgc cagactggcc	480
cacgacatga tcgtccgcga cccgggcctg tcgtcgcgc aggtcgcgc cgccgccggg	540
atctcgcgc agaccgtcag ggacgtcaga caccggatgc tccgcgtga ggaccggtg	600
ccgcgcgcgc ggccgcggac cctggtggag cgcggcgcgc accgccgggc ggagccggcc	660
gggaaggccg ccgcgccgtg cgggacggag ccgccgccgc ccgtcgtgat gaagcggctg	720
agggccgac cggcgtgcg tctcaacgag aacggacgc acctgctgc gcttctggat	780
atccacacgg tccggtgga ggactggaac cgcattatcg aaagcgtgcc gccgcaccgt	840
ctggagacgg tggcgagct ggacgctcc tgcgccgaca aatggtccga gatcgcgtca	900
cgcacgaaa gcaacgcac acatctggcc ggtgta	936

<210> 39

<211> 39331

<212> ADN

<213> Secuencia Artificial

<220>

<223> fosmid pXYF148 Sint 騁 ico

<400> 39

atcgaatata acttcgtata atgtatgcta tacgaagtta ttagcgatga gctcggactt	60
ccattgttca ttccacggac aaaaacagag aaaggaaacg acagaggcca aaaagctcgc	120
tttcagcacc tgtcgtttcc tttcttttca gagggtatit taaataaaaa cattaagtta	180
tgacgaagaa gaacggaaac gccttaaacc ggaaaatitc cataaatagc gaaaaccgcg	240
gaggtcgccg ccccgttaacc tgtcggatca ccggaagga cccgtaaagt gataatgatt	300
atcatctaca tatcacaacg tgcgtggagg ccatcaaacc acgtcaaata atcaattatg	360
acgcaggtat cgtattaatt gatctgcatc aacttaacgt aaaaacaact tcagacaata	420
caaatcagcg aactgaata cggggcaacc tcatgtccga gctcgcgagc tcgtcgacag	480
cgacacactt gcatcggatg cagcccggtt aacgtgccg cagggcctgg gtaaccaggt	540
atitgtcca cataaccgtg cgcaaatgt tgtggataag caggacacag cagcaatcca	600
cagcaggcat acaaccgcac accgaggtta ctccgttcta caggttacga cgacatgtca	660
atacttgccc ttgacaggca ttgatggaat cgtagtctca cgctgatagt ctgatcgaca	720
atacaagtgg gaccgtggtc ccagaccgat aatcagaccg acaacacgag tgggatcgtg	780
gtcccagact aataatcaga ccgacgatac gagtgggacc gtggtcccag actaataatc	840
agaccgacga tacgagtggg accgtggttc cagactaata atcagaccga cgatacgagt	900
gggaccgtgg tcccagacta ataatcagac cgacgatac agtgggacca tgggtccaga	960
ctaataatca gaccgacgat acgagtggga ccgtgggtccc agtctgatta tcagaccgac	1020
gatacgagtg ggaccgtggt cccagactaa taatcagacc gacgatacga gtgggaccgt	1080
ggtcccagac taataatcag accgacgata cgagtgggac cgtgggtcca gtctgattat	1140
cagaccgacg atacaagtgg aacagtgggc ccagagagaa tattcaggcc agttatgctt	1200
tctggcctgt aacaaaggac attaagtaaa gacagataaa cgtagactaa aacgtggtcg	1260
catcagggtg ctggcttttc aagttcctta agaatggcct caattttctc tataactca	1320

gttgaacac gagacctgtc caggttaagc accattttat cgcccttata caatactgtc	1380
gctccaggag caaactgatg tcgtgagctt aaactagttc ttgatgcaga tgacgtttta	1440
agcacagaag ttaaaagagt gataacttct tcagcttcaa atatcacccc agcttttttc	1500
tgctcatgaa ggttagatgc ctgctgctta agtaattcct ctttatctgt aaaggctttt	1560
tgaagtgcac cacctgaccg ggcagatagt tcaccggggt gagaaaaaag agcaacaact	1620
gatttaggca atttgcggtt gttgatacag cgggtaataa tcttacgtga aatatittcc	1680
gcatcagcca ggcagaaat atttcagca aattcattct gcaatoggct tgcataacgc	1740
tgaccacgtt cataagcact tgttggcgga taatcgttac ccaatctgga taatgcagcc	1800
atctgctcat catccagctc gccaacaga acacgataat cactttcggg aagtgcagca	1860
gctttacgac ggcgactccc atcggcaatt tctatgacac cagatactct tcgaccgaac	1920
gccggtgtct gttgaccagt cagtagaaaa gaagggatga gatcatccag tgcgtcctca	1980
gtaagcagct cctggtcacg ttcattacct gaccataccc gagaggtctt ctcaacacta	2040
tcaccccgga gcacttcaag agtaaacctc acatcccagc cacatacagg caaagtaatg	2100
gcattaccgc gagccattac tcctacgcgc gcaattaacg aatccaccat cggggcagct	2160
gggtgtcgata acgaagtatc ttcaaccggt tgagtattga gcgtatgttt tggaataaca	2220
ggcgcacgct tcattatcta atctcccagc gtggtttaat cagacgatcg aaaatttcat	2280
tgcagacagg ttcccaaata gaaagagcat ttctccaggc accagttgaa gagcgttgat	2340
caatggcctg ttcaaaaaca gttctcatcc ggatctgacc tttaccaact tcatccgttt	2400
cacgtacaac attttttaga accatgcttc cccaggcatc ccgaatttgc tcctccatcc	2460
acggggactg agagccatta ctattgctgt atttggttaag caaaatacgt acatcaggct	2520
cgaacccttt aagatcaacg ttcttgagca gatcacgaag catatcgaaa aactgcagtg	2580
cggaggtgta gtcaaacac tcagcaggcg tgggaacaat cagcacatca gcagcacata	2640
cgacattaat cgtgccgata cccaggttag ggcgctgtc aataactatg acatcatagt	2700

catgagcaac agtttcaatg gccagtcgga gcatcagggtg tggatcgggtg ggcagttttac	2760
cttcatcaaa tttgcccatt aactcagttt caatacgggtg cagagocaga caggaaggaa	2820
taatgtcaag ccccgccag caagtgggct ttattgcata agtgacatcg tccttttccc	2880
caagatagaa aggcaggaga gtgtcttctg catgaatatg aagatctggt acccatccgt	2940
gatacattga ggctgttccc tgggggtcgt taccttcac gagcaaaaca cgtagcccct	3000
tcagagccag atcctgagca agatgaacag aaactgaggt tttgtaaag ccacctttat	3060
gggcagcaac cccgatcacc ggtggaata cgtcttcagc acgtcgcaat cgcgtaccaa	3120
acacatcacg catatgatta atttgttcaa ttgtataacc aacacgttgc tcaaccgctc	3180
ctogaatttc catatccggg tgcggtagtc gccctgcttt ctcgcatct ctgatagcct	3240
gagaagaaac cccaactaaa tccgctgctt cacctattct ccagcgccgg gttattttcc	3300
tgccttcggg gctgtcatca ttaaaactgtg caatggcgat agccttcgtc atttcatgac	3360
cagcgtttat gcactggtta agtgtttcca tgagtttcat tctgaacatc ctttaatcat	3420
tgctttgcgt ttttttatta aatcttgcaa ttactgcaa agcaacaaca aaatcgcaaa	3480
gtcatcaaaa aaccgcaaag ttgtttaaaa taagagcaac actacaaaag gagataagaa	3540
gagcacatac ctcagtcact tattatcact agcgctcgcc gcagccgtgt aaccgagcat	3600
agcgagcgaa ctggcgagga agcaaagaag aactgttctg tcagatagct cttacgtca	3660
gcgcaagaag aaatatccac cgtgggaaaa actccaggta gaggtacaca cgcggatagc	3720
caattcagag taataaactg tgataatcaa cctcatcaa tgatgacgaa ctaacccccg	3780
atatcaggtc acatgacgaa gggaaagaga aggaaatcaa ctgtgacaaa ctgccctcaa	3840
atttggcttc cttaaaaatt acagttcaaa aagtatgaga aaatccatgc aggctgaagg	3900
aaacagcaaa actgtgacaa attaccctca gtaggtcaga acaaatgtga cgaaccacc	3960
tcaaattctgt gacagataac cctcagacta tcctgtcgtc atggaagtga tatcgcgga	4020
ggaaaatacg atatgagtcg tctggcgcc tttctttttc tcaatgtatg agaggcgcat	4080
tggagttctg ctgttgatct cattaacaca gacctgcagg aagcggcggc ggaagtcagg	4140

catacgctgg taactttgag gcagctggta acgctctatg atccagtcga ttttcagaga	4200
gacgatgcct gagccatccg gcttacgata ctgacacagg gattcgtata aacgcatggc	4260
atacggattg gtgatttctt ttgtttcact aagccgaaac tgcgtaaacc ggttctgtaa	4320
cccgataaag aagggaatga gatatgggtt gatatgtaca ctgtaaagcc ctctggatgg	4380
actgtgcgca cgtttgataa accaaggaaa agattcatag cctttttcat cgccggcatc	4440
ctcttcaggg cgataaaaaa ccacttcctt ccccgcgaaa ctcttcaatg cctgccgtat	4500
atccttactg gcttccgcag aggtcaatcc gaatatttca gcatatttag caacatggat	4560
ctcgcagata ccgtcatgtt cctgtagggt gccatcagat tttctgatct ggtcaacgaa	4620
cagatacagc atacgttttt gatcccgga gagactatat gccgcctcag tgaggtcgtt	4680
tgactggacg attcgcgggc tatttttacg tttcttgtga ttgataaccg ctgtttccgc	4740
catgacagat ccatgtgaag tgtgacaagt ttttagattg tcacactaaa taaaaaagag	4800
tcaataagca gggataactt tgtgaaaaa cagcttcttc tgagggaat ttgtcacagg	4860
gttaaggga atttgtcaca gacaggactg tcatttgagg gtgatttgtc aactgaaaag	4920
ggcaatttgt cacaacacct tctctagaac cagcatggat aaaggcctac aaggcgctct	4980
aaaaagaag atctaaaaac tataaaaaaa ataattataa aaatatcccc gtggataagt	5040
ggataacccc aagggaagtt ttttcaggca tcgtgtgtaa gcagaatata taagtgtgt	5100
tccctgggtc ttcctcgctc actcgaccgg gagggttcga gaagggggg cccccctt	5160
cggcgtgcgc ggtcacgcgc acagggcgca gccctggta aaaacaaggt ttataaatat	5220
tggtttaaaa gcaggttaaa agacaggtta gcggtggccg aaaaacgggc ggaaaccctt	5280
gcaaatgctg gattttctgc ctgtggacag ccctcaaatt gtcaataggt gcgcccctca	5340
tctgtcagca ctctgccct caagtgtcaa ggatcgcgcc cctcatctgt cagtagtcgc	5400
gccctcaag tgtcaatacc gcagggcact tatcccagg cttgtccaca tcatctgtgg	5460
gaaactcgcg taaaatcagg cgttttcgcc gatttgcgag gctggccagc tccacgtcgc	5520

cggccgaaat cgagcctgcc cctcatctgt caacgccgcg ccgggtgagt cggcccctca	5580
agtgtcaacg tccgcccctc atctgtcagt gagggccaag tttccgcga ggtatccaca	5640
acgccggcgg ccggccgcgg tgtctcgcac acggcttoga cggcgtttct ggcgcgtttg	5700
cagggccata gacggccgcc agcccagcgg cgagggcaac cagccgaggg cttcgccctg	5760
tgcctcgact gcggcgagca ctactggctg taaaaggaca gaccacatca tggttctgtg	5820
ttcattaggt tgttctgtcc attgtgaca taatccgctc cacttcaacg taacaccgca	5880
cgaagatttc tattgttctt gaaggcatat tcaaatcggt ttcgttaccg cttgcaggca	5940
tcatgacaga acactacttc ctataaacgc tacacaggct cctgagatta ataatgcgga	6000
tctctacgat aatgggagat tttcccgact gtttcgttgc cttctcagtg gataacagcc	6060
agcttctctg tttaacagac aaaaacagca tatccactca gttccacatt tccatataaa	6120
ggccaaggca tttattctca ggataattgt ttcagcatcg caaccgcctc agactccggc	6180
atcgcaaaact gcacccggtg ccgggcagcc acatccagcg caaaaacctt cgtgtagact	6240
tccgttgaac tgatggactt atgtccctc aggctttgca gaactttcag cgggtataccg	6300
gcatacagca tgtgcatcgc ataggaatgg cggaacgtat gtggtgtgac cggaacagag	6360
aacgtcacac cgtcagcagc agcggcggca accgcctccc caatccaggt cctgaccgtt	6420
ctgtccgtca cttcccagat ccgcgctttc tctgtccttc ctgtgcgacg gttacgccgc	6480
tccatgagct tatcgcaat aaatacctgt gacggaagat cacttcgcag aataaataaa	6540
tcctggtgtc cctgttgata ccgggaagcc ctgggccaac ttttgcgaa aatgagacgt	6600
tgatcggcac gtaagagggt ccaactttca ccataatgaa ataagatcac taccgggcgt	6660
atTTTTTgag ttatcgagat tttcaggagc taaggaagct aaaatggaga aaaaaatcac	6720
tggatatacc accgttgata tatcccaatg gcctcgtaaa gaacattttg aggcatTTTca	6780
gtcagttgct caatgtacct ataaccagac cgttcagctg gatattacgg cTTTTTaaa	6840
gaccgtaaag aaaaataagc acaagTTTT tccggccttt attcacattc ttgccgcct	6900
gatgaatgct catccggaat ttcgtatggc aatgaaagac ggtgagctgg tgatatggga	6960

tagtgttcac ccttggtaca ccgttttcca tgagcaaact gaaacgtttt catogctctg	7020
gagtgaatac cacgacgatt tccggcagtt tctacacata tattcgcaag atgtggcgtg	7080
ttacggtgaa aacctggcct atttccctaa agggtttatt gagaatatgt ttttctctc	7140
agccaatccc tgggtgagtt tcaccagttt tgatttaaac gtggccaata tggacaactt	7200
cttcgcccc gttttcacca tgggcaaata ttatacga ggcgacaagg tgctgatgcc	7260
gctggcgatt caggttcatt atgccgtttg tgatggcttc catgtcggca gaatgcttaa	7320
tgaattacaa cagtactgcg atgagtggca gggcggggcg taatTTTTT aaggcagtta	7380
ttggtgccct taaacgcctg gttgctacgc ctgaataagt gataataagc ggatgaatgg	7440
cagaaattcg atgataagct gtcaaacatg agaattggtc gacggcccgg gcggccgcaa	7500
ggggttcgcg ttggccgatt cattaatgca gctggcacga caggtttccc gactggaaag	7560
cgggcagtga gcgcaacgca attaatgtga gttagctcac tcattaggca ccccaggctt	7620
tacactttat gcttcggct cgtatgttgt gtggaattgt gagcggataa caatttcaca	7680
caggaaacag ctatgaccat gattacgcca agctatttag gtgagactat agaatactca	7740
agcttgcatt cctgcaggtc gactctagag gatccacca ccgccggacg ggggtcggcg	7800
gcccgaagt ccaccaggat cgagccggcg atctgccgcc cgcgccacca cgactccatc	7860
agcccgacct cctcctccgg gctgcgcacc agccgcagca gcagcgagca ggaccgctcc	7920
accaggacgc tctccacccc ggagatgaac tccatgtaga acggctccag gccgagcagc	7980
cgggcggggc ggagaccgc gagaccacc acgtccacc cgcaccggc cagcgtgcgc	8040
gcggttcggc tcggcgccca cccagctcc cgcgccgcc ggaagatgcg gtcccggtc	8100
gcctccgaca gcccgggctt ccggttgaag gcgagggaca cggcgccctt ggacacgccg	8160
gcgcgcgcgg cgacgtccct gatggtgac cgaggggtcg gcgttgccgt catcgagtgg	8220
gtccacgca gtacaggcg gaacgggcgg tgtccgggtc cggagtcttc cccccgca	8280
ccccgatggt cacctgttcc ccggggagca gggtcaccag ccccggtcg gccgcgccc	8340

cggggtccag ccggtcggcc tggagcagca ggtcccgtac gaggtgctgg gccgtgaccg	8400
tgatcccgtc cggcgcgagg gcgacctcga actccggcgg ggggtagggg atctcccgtt	8460
ccggcgccgg gaagtgccac gcccgacccc cgtccgcgtc ggcgaccagg aactccccgg	8520
ggccgtccgg cagcagttcg accgggacct cgaccacggc caccgtccgc cccccggcgt	8580
ccagcgccgg gccgcctcc gcgatcgggg cgcgtcgac ggacatccgg cgcagccgca	8640
gcgttccccg ccagtctctc gcggactggt tgaccgccgc caccaccaga ccgtcaccgt	8700
ccgcgcgcac ggtcagcagc cggtcgcgt acagccggcg cagctcgtgg tagagcggct	8760
tctccgccc gtccccgtc atcgcggccc acgacgtcac cggccagcag tcgttgagct	8820
gccagaccac cgtgcccgcg cacaccggcc agtgcgagcg ccagtgtctg acaccggccg	8880
ccaccgcacg cgcctggttg acctgcgtca gatagtcca gcggtcgaag tcgccctccg	8940
gcacggcgaa gtggcgggcg aggccgcgct ccagcttgcc gttgccgtcc tccgccttct	9000
ggtggtgcag catgccgggg gagtccggcg cggggtcctc cccgggcagc gccgcgcgca	9060
gcgtggcgtg cgcgggaggc gcctgccagc cgaactcggc cacgaagcg cggacgtcgc	9120
gccggtagtc ggcgtagtcg gcgcggttcc acacctcca ggagtgggtg gtgccgtgcg	9180
ccgatcgtt ggggtggtgc gccaggaac cggaccaggg actgcccgc gtgtacggcc	9240
gcgtcgggtc cagctccgcg accacccgcg gcaggacgcc gaggtagtag ccctcgcccc	9300
aggagtcccc ggcgagcccc tgctcccagt ccagctccg gaacccccac aggttctcgt	9360
tgttgccgtt ccacagcacc agggaggggt gcggcatcag ccgtacgacg ttctcccggg	9420
cctccgcctc cacctccccg cgcagcggct gtcctcggg gtaggcggcg cacgcgaacg	9480
ggaagtcctg ccagaccagc agccccaact cgtcgcaggc gtcgtagaag tcctcgtcct	9540
cgtagatccc gccgccccag acccggacca ggtccacccc cgcgccggcc gcctgctcca	9600
gccggtgccg gtagcgtcc cgggtgatcc gggacgggaa cacgtcgtcc gggatccagt	9660
tgacgccccg cgcgaacagc cgctcaccgt tgacgaccag ggtgaacccg gtgccgtgcg	9720
cgtcggccga ggtgtccagc tcaaccgtcc ggaaccgggt cctgcgccgc caggcgtcca	9780

gagcctcgtc accgtgggac aacgtcagct cgacgtcgta cagcggctgt tcgccgtatc	9840
cgcgcggcca ccacaggcgg acgtccggca cccggagccg cacggtcccg gccgtcccat	9900
cgacccgcgc cggggcgcgc acgccccggg cgctcgcctc cagggtgagc ggtgcctcga	9960
cccgggagcg ctccacgtcg accgccagct cgatctgccc caccocgtcc tcgacggtga	10020
ccagcgggcg caccggggcg atccgcgccg tcgaccagcg ctccagccgc accggccgcc	10080
agatcccggc cgtcaccagc gtccggcccc agtcccagcc gaacgagcag gccatcttcc	10140
gcaggtactg gtacggctcg gcgtacgctc cggggcgctc gccagcctg ccgcgcaccg	10200
cctccgcctc ggcgtaacgc gaggcgaacc gcaccgtgag ccggccgctc agtcccgtca	10260
cgtcgaagcg gtacgagcgg tgcattgttc gcgtccggcc cagtggccgg ccgtcgagca	10320
ggatctcggc gacggtgtcg agaccgtcga agacgaggtc cgtctgctcg tgcgggcccc	10380
tcccggcggc cagctccgtc tcgtacgtcc actcccgcgc gccaccag gccacctcgg	10440
tctcgttgcg gccgaggaac ggatcgggga tcagcccggc gccagcaga tcggtgtgca	10500
cacacccggc caccgaggcg gggaggcggt ccccgctgcc gtccgggtgt cgcaggatcc	10560
atccctcggc gagcgggtgt acctgacgca tgcacactcc ctaaaccggt tgagccttct	10620
ctgaagagtg gtctggcctc gttggcgca ttgcacttt accggttcag ttcagggtg	10680
ccagagtgcc gaatcagcca tccactcgt gctcgtccgt cccgtgaacg gagccgtgat	10740
gcatttgaac cgcctacga cactcaccgg atcgtcgc ctgctgcgc tcctggcctc	10800
cgcctgcacg ggcacggggg gttcctcga gggcgcgac gccaaaggctc ccgacgacc	10860
gtcaaaggctc aagggtccc tcacggtcct caccaccgg accgatctgg tgcaggacgg	10920
gacgatgaag aagtacgcc cagagttcaa cgagacctat cccggggtga aggtggagtt	10980
cgacggcctc accgactacg agggcgaggt caagatccgt atgaacacgg agaactacgg	11040
cgacgtcctc atgatccgg cggctcgtcga gaagaaggac taccgaagt tcttcgcctc	11100
cctgggcacc aaggccgaac gcgcgcgcaa gtaccggttc accgactact ccaccgtcga	11160

cggcaaggtc tacgggcaga gccccgtcgg cgtcgtcccc gggttcattc acaacaagcg	11220
ggtgtggagc gagggccggc tcaccgactg gccaccacc cccgccgagt tcctggacga	11280
cctgaaggcg atccggtcga agaccgacgc ggtgccgtac tacaccaact tcaaggacat	11340
gtggccgctg acccagtgga ccaacgtcaa cggctccgtc ggctgcgacc cgcacgccac	11400
cacgaagctc gccgagggcg acccgtgggc cgagggggcc gacctgcgcg tgggcgacac	11460
cctgctccac gacatcgtgc gcggcggact cgcgcagaag gacccgacca ccaccaactg	11520
ggagggctcc aagcccaagc tggccaaggg cgagatcgcc accatgtggc tgggctcctg	11580
ggccgtcgtg cagatcgggg acgcggcgaa gcaggccggc gccgaccccg ccgacatcgg	11640
cttcattgcc ttccccgcac agcgggacgg cacgtttctg gcggtgacct ccccgacta	11700
ccagcaggcg gtcaacgtca actccgacaa caaggaggcc gcccgccctt ggatcgactg	11760
gttcaccgac aagtccggct acgccgaggc caacctcgcc ctatcccccc tgaaggacgc	11820
cccgtgccc gccgtcctcg agccctacga gaaggccggc gtgaagctcc tggacctcga	11880
ggacagcaag ggccgccagg tgaagtccct cgacaaccgc tccgaggtcg gcatctacaa	11940
gcccgactac cgccaggaac tcgtcgacct cgcgcgggc gcccgcaagg gcggcctgga	12000
cgactacctc ggccggcctcg gcgagcgctg ggccgaggcg cgcagcgcg cgggggcctg	12060
atgacggaca ccaccgcaa ggccggcgcg ccggttcccc cggccgcgcc gcgccggccg	12120
ggccgcgcgg ccgccgcccc gcgccgcacc cggctgtcgc gccgcctcac cccgtggctg	12180
ttcctggccg caccgtggc cctgtcctg accttcacct acgcgccga tcgccaacat	12240
ggtcgcgtac agcttcaccg actgggacgg cgtgagcccg gagctgaact ggacggcac	12300
cgggaactac accgaactcc tcaccgctc cgagctgttc gaggtcttct tcgtcagcg	12360
ctactacctc gtgcctccg cgggtcagat cgtgtcgcct ctctacttcg ccacggtcct	12420
cagcttcgac gtccgcttc ggaacttctt caaggcggtg ctgttcttc cgtacctcat	12480
caacggggtg gccatcggct tcgtcttct ctacttcttc caggacggcg gcaccctcga	12540
ctccgtactg ggccgtcctg gcgtcgagac cgaccacgcc tggctgggca cgcggttctc	12600

cgogaacacc tcgctggccg gcgtctccgt ctggcgctac ctcggaactga acttcgtcct	12660
cttcctcggc gcgatccagt ccatcccggg cgagctgtac gaggcggccg agatcgacgg	12720
cggaaccgc tggcagcagt tccggcacat catcgcgccc ggcatcagac ccgtgctgag	12780
cctgagcgtg atcctctcgg tctccggctc gctgtcggtc ttcgagatcc cgtacatcat	12840
gaccggcggc gccaccggca cggagacctt cgtgatccag accgtgaagc tggcgttcca	12900
gttcaacaag acgggactcg cctcggccgc cgccgtcgtc ctgctgctga tcgtcctggc	12960
ggtcacctgg gtgcagcggc gcatcgtccc cgacgagaag gtggacctcg tatgaccgc	13020
cgtaccgcgg cacgcgccct ggtcctgacg tccctgatcc tggcgacgct ggtggtgctg	13080
ctgccgctcg ccgtggtctt cctgacctcg ctgaagtcct ccgaggagat ggogaacggc	13140
agcggagcgc tgaccgcccc cgacgaccgc ctgaacttcg gcaactacgt gacggcgttc	13200
cgggacggcc agatgctgtc cgcgttcggg aacacggccg tcatcctggt cgtggccgtc	13260
ggcggaacga tctgatcgg ctgatgacg gcgtacgca tcgaccgctt ccggttcgc	13320
ttcaagaagc tggctgtggc gctgttcctg ctggccgcgc tggccccgg ggtgaccacc	13380
caggtggcga ccttcagat cgtcaacagc ttccgcatgt tcgacagcct gtggcgccg	13440
atgccctct acatgggcac ggacatcgtc tcgatctacg tcttcctgca gttcatccgc	13500
tccatccccg tctccctgga cgaggcggcg cgcctggaag gcgccaacgc gttcaccgtc	13560
taccgaagg tgatcttccc gctgctcaag ccggcgatcg cgacggtggt gatcgtaaag	13620
gggatcaacg tctacaacga cttctacatc cccttcctct acatgccctc cgaggacctg	13680
ggggtcatct cgacgtccct gttccgcttc aaggccccct tcggcgcgca ctgggagacg	13740
atctcggcgg gcgcggtcct ggtcatcctg ccacacttga tcgtcttcct gttcctccag	13800
cgttcctct acaacgggtt catgcggggg gcgacgaagt agccagcgc gccaccagca	13860
cgggagtac ctggtcgacc tgccacgcgc gtgccccgtg cgccgtcagc gccgcggcga	13920
cggaccgctc gtcgggcccc ggccggtccc agcagacctt gcgcaccgtg tccggggtga	13980

tcaggttctc cggcggcatg ttcagccgct cggcgagttc ggcgaccccc gcgcggggccg	14040
cgcacagccg ggccgcggca acggggtcct tgtccgccca ggcgcgcggc ggcgaggggc	14100
cggtcaccgg ctggccgggc tgcggcagct gggcctcgct cagcgccctc gcgcggtcga	14160
cggccgcctg ccactgctcc agctggcgcc gccccaccgc ctgcccgaac ccgttgagcg	14220
cggccatggc gtgcaggttg gcgggcagcg cgagcgcggc ctccacgac gccgcgtcgg	14280
aaagcacctt gccgggggag acgtcacggc gccggcgat ccggtcgcg gtctcccaca	14340
gtccccgcac caccgccatc tggcggcgcc ggcgcacctt gtgcatgcc gaggtgcggc	14400
gccaggggtc cttgcggggc tccggcgcg gggccgaggc gatcgcgctg aactcctgcc	14460
gggcccagtc cagcttgccc tggcgggtcca gtccttctc caggcggtcc cgcagatcga	14520
ccagcagttc gacgtcgagg gcggcggtacc gcagccaggc ctcgggcagc ggacgggtgg	14580
accagtcgac ggcgggagtgg cccttctcca ggacgaagcc gagcacgttc tcgaccatcg	14640
cgccgagccc gacgcggggg aaccgggcaa ggcggccggc cagctcggtg tcgaagaggc	14700
gggagggcac catgcctatc tcgcgcagac agggcaggtc ctgggtggcg gcgtgcagca	14760
cccactcgac gccggacagc gcctcgccga gggcggacag gtcggggcag gccacggggt	14820
cgatcagcgc ggtaccgcga ccctcgcggc gcagctggac gaggtaggcg cgctggccgt	14880
agcggtagccc ggagcgcgcc tcggcggtcca cggcgacggg tccgctgcc gccgcgaagg	14940
cggcgaccgc ctcggcgagg gcggcctcgt ccgctatcac gggcggaatg ccctcgcggg	15000
gttccagcaa gggggtcggc gcctccgtaa cagaagatcc gccgtcgtcc ggaggagcgc	15060
ctccggtggt gcgcagtga ctgtctgctg cgggtgctg ggctcggtc acctgtcaag	15120
ggtatccgtg ccgcgaaggc gcccgctgac ggttgtgctc cgtgacgggc gccggtgggt	15180
cgtattccgg tcagaagagt gaaagaacgt gttcgcttgg ccgtgggcgg gcggatcggg	15240
gacgggcgga tcgggggcgg gacggaaggc tcagtggatg atgccggtgc gcagggccac	15300
ggccaccatg ccggcgcggt cggccgtgcc gagcttgagg gcgatccgg ccaggtggct	15360
cttgacggtc agtgcggaca gggccatcga gacgccgac gccttggtcg actggccttc	15420

cgccaccagc cgcagcacct ccacctcgcg gccggacagc tcgcggtagc cgcccgggtg	15480
gctcggggca cccggggggc ggcggtgcag gcgcgcggcg gcggcgccga tgggagcggc	15540
acccggccgg gtggggagcc cgaggttggg gcgggtgccg gtgacgacgt agcccttgac	15600
tccgcccgcg agggcgttgc gcacggcgcc gatgtcgtcc gccgcggaga gggcgaggcc	15660
gttgggccag cccgcggcgc ggtctcggg gaggagggtg aggccggaac catccggcag	15720
atggacttcg gcgacgcaga tgtcgcgggg gttgccgatg cggggacgag cctccgcgac	15780
ggacgaggcc tcgatgacgt cgcgtacgcc gagcgccac aggtggcggg tgacggtcga	15840
acggacgcgg gggtcggcca cgaccaccat ggcggtcggc ttgttcgggc ggtaggcgac	15900
caggcttgcg ggctgctcga ggagaacgga caccaggcct cctgggggtgc gggacgggcc	15960
ggctcgtggg ggtgaaggcg ggacgaaccg tgctttcaag gtcacagtcg tcttcggcag	16020
caaacctggt gtcctttaac gaatgatcac gaagtgatga gtaacaatcc gggcaattcg	16080
gagcacgat cgatcattcg aagatcgaac ggtttcggtc tgcgtcgcaa cgcttcgaa	16140
agtggccgta tcgacaaaga gagatgcagg aggcgggtcg tcgggacccc gcagcgggag	16200
gctcagcgcg actgcggccc cctccgctgc ggcagcgta ccacggacgc gtccccgga	16260
gcggccggcg gcagccccgc gacctgcgcc agcagatcgg accacgcgac cagatgggcg	16320
gccgtgtccg gaaccccgcc cagaccctca cgcggcgtcc acgaggcacg gatctcgatc	16380
tgggaggcgg cgggccgcgc ggacagccc cccaagtagt gcgaactcgc ccgcgtcacc	16440
gtccgctcg gctcgccgta cgacaggccg cgcgcccca gcgcgccgt cagccaggac	16500
cagcacacgt ccggcagcag cggatccgcc gccatctccg gctccagctc ggcgcgacc	16560
agcgtacca gacggaaggt cccccgccag gcgtcgtgtc cggccgggtc gcacagcagc	16620
accagccggc cgtcggccag atcctcctcg ccgtcgacga ccgccgcctc cagcgcgtgc	16680
gcgtacgggg cgagccgttt cggcgcgggc accgtctcca cctcgatctg cggccgcagc	16740
cgggcgctct gcagcgctc gacagcggcc cggaaggcg gcggaggcgc acctccgtgc	16800

ccgggatccc ccccgccctc cttcggttcg tccattccgc cagcgccgtc cgacagtcgt	16860
ccctgagccg cagccatgcc gggaagatta agcggaacgg gcccccgcg cagggaggga	16920
caccgcgcc gcccgcgct gtccggatcc tgcaccggg ccccgccgc cggacgcccc	16980
tggggtccgg ggcgcgggac gggtcgtgcg agactggccg gtgtgagtgc caacacgagc	17040
ccgaagggcc agacgcctac cgcgaccccc gaccccgta agaacgacgc cgtccgggaa	17100
tcagccttcc tcaaggcgtg ccggcgcgag ccggtgccgc acacgccgt gtggttcatg	17160
cggcaggccg ggcgctcact gccggagtac cgcaaggtag gcgaggcat cgggatgctc	17220
gactcctgca tgcggccga gctggtcacc gagatcacc tccagccgt gcgcccac	17280
cacgtcgacg cggcgatcta cttcagcgac atcgtcgtcc cgctcaaggc catcggcac	17340
gacctcgaca tcaagccgg catcgcccg gtcgtcgagc agccggtgc caccgcgcc	17400
gacctcgccc ggctgcgga cctgaccccg gaggacgtct cctacgtcac cgaggccatc	17460
ggcatgctga cccgtgagct cgggtccacc ccgctgatcg gtttcgagg cgccccgttc	17520
acccttgca gttacctgt cgaggcggg ccgtccgta cgtacgagaa cgccaaggcg	17580
atgatgtacg gcgacccga gctctgggcc gacctgctcg accgcctcgc cgacatcacg	17640
gggccttcc tcgacgtcca gatccgggcc ggcgctcgg ccgtgcagct cttcgactcc	17700
tgggcggcg cgctcgccc ctcgactac cggcgttcgg tgctgcccgc ctggcggaag	17760
gtgttcgcg cgggtggcgg ccacggcgtc ccgcgcatcc acttcggcgt cggcaccggc	17820
gagctgctgg ggctcatggg cgaggccggc gcggacatcg tcggcgtcga ctggcgctc	17880
ccgatggacg aggcgcccc ccgcgtcggc cccggcaagg cgctccagg caacctggac	17940
ccgaccgtgc tgttcgccc ccgggaggcc gtcgagacga aggcgcgga ggtcctggac	18000
accgcccgg gcctggagg ccacatcttc aacctcggtc acggagtgat gccctccacc	18060
gaccggacg ccctcaccg tctcgtggag tacgtccaca cgcagacggc gcgctgacct	18120
accgctcacg cgccggacgc gagtccgaat ccggggcggg ggtactgggc acgggtgccc	18180
accacgttca ccccgggta cgggcaggtag gaggcccat gaggctcgag atgttcgacc	18240

cgcgccgat cggcgtcgtg ttcaccagg ggccggagca ccggctcgcg tacaccaacg	18300
ccgtctaccg ggagaccttc ggcgaccgcc cgctggggcg gacgatccgc gaggccttcc	18360
ccgacctcgc gcagtcgggc tacttcgaca tcttcgaccg ggtcctcacc acgggcgcgg	18420
ccgaggtggt caccgcggtg cccctcgacc tgatctaccc ccgctccacg ggcgagggca	18480
ggcgctactt cacgttcagc atctcccgcg ccacgatgag cgacggccgg ccgggagtg	18540
tggcgtgat cgtggagggt accgcgcagg tgaccgccgc ggaacggatc cgtgtgctgg	18600
ccgaggagcg ccgccgcgcg ctgcagcgct accgcagcct ggtgaacgcc ggaacgcaga	18660
tggtgtgggt ggcggaagcc aaggggccga tcaccgagcc gagccccggc tgggaacgcg	18720
tgaccgggca gacctgggag gagttccgcg gcgagggtg gatgaacgcc gtccaccccg	18780
acgaccgcgc gcctcggtc gaggcgtggc ggcgggcgac gaccgaacag gtgccgcgct	18840
ggatccacac ctaccggctg ccgctggccg ccggcgggta ccggcacttc gtctcgacg	18900
ccgcgccgt gcgcgaaggg aacacggtga tcgaatgggt gggcacctgc acggacatcg	18960
agcgggaatg gcaggagggc gcccgtagg aactgctggc gcgggcccgc accgccacgt	19020
ccggcatcgc gcggctggac gagatgctcg ccgccctggc cgatgtgatc gtgcccgaca	19080
tgcgcgaaa ctgcaccatc cacctcctgc cgcaggccct gcaccgtctg ccgggcaccc	19140
cgctgaccac cgaacgcgtc gccgcggtca cccgccggg gctcccggac ctgccccgc	19200
accacgagga gcacctgcgg cccggcagcc cgctggcccgc gccgcgcgac cgccgcagcc	19260
cgctccactt cgtcttcccgc cccggcgagc cgccggccga cctcgtccg ctogacggcg	19320
agccctggat ggccgaggac gtcaacagcg tcgtgctgct gcccgctcgc gtogacggca	19380
ccaccgccgc cctggtcgcc gtctccacca gcggcgcccgc cccgcccctc ggccaggcgg	19440
agatcggcct gctgcagaca ctctggaac gcgcccacac cccctcagc aacgccctgg	19500
agtaccagcg caccggcag gtggccctgg ccctgcagaa cagcctgctc accgaccgc	19560
cggacgcgcc cggcctggac atcgccgtcc gctaccggcc cagcaccgcc gccgccagg	19620

tcggcgggga ctggtacgac gcgttcgtgc tgcgcgacgg cgccaccgtc ctcaccatcg	19680
gcgacgtctc cggccacgac ctgocggccg ccgtcaccat gagccagctg cgcaacatgc	19740
tgcgcgggct cacgctggac cgccaggaac cgaccggcac catcctgcgc cggctggaca	19800
tcgccgtgca gacctctat acggagtgca ccgccacctg cgtgctggcc cgggtggaac	19860
gcccggactc cggcggcgtc cggctgcact actccgtcgc cggtcacccg ccgccgctgc	19920
tcgtcgaggc ggacggctcc gcgcgcttcc tgaccggggc gcggtcccg atgctcgggc	19980
tcgtcccgcc gccggagtac tcgagcgcca tggaaccgt gccgcccgcc tccaccctgc	20040
tgctgtacac cgacgggctg gtggagcgcc gcgacgagga tctcaccgtg ggcttgagc	20100
ggctgcggca ccacgcctcg gaggcggtca gccgcccgct gcaggacttc tgcgacacac	20160
tgctcaccgg ccagctcacc gtcgacaacg acgacgacgt ggcatgctg gtctgcgcc	20220
ggtaggagcg tgccgaggag cgccactctg gccgatttta cccttgcttt tccatcggga	20280
ttcgttctcc ggatttccg atccggcgcc gacggcgaga ccgttgggat caccaatacc	20340
ccggaattcc cgctccgcc accgttgggc agcgacggat cctgtgatat ttogactacg	20400
cgcggtgatg aattggctcg gtgccggtcg cggccggtg tagcagttct ggagcgcgtc	20460
tggacatcgt cacgagcgct tgtgattctt ggtcctgtac acgcaagccg gcgcaacgtc	20520
cacgttgccc atcagcgggt atcggcggtc caccggcgcg acggtgaccg cgggcgggta	20580
ctcatagggg gaactgcaat gaattactca aaagcagcga gaggaatgcc gacagccgga	20640
caaggtgccg ttcgggcggc gcgcgtcgtc cgtgaaagtc cggcggaatc agaaacggtc	20700
acagttcaga tagcgtcgtt attacgggt gagtcgctgc gctcgaaagg gatcgagcag	20760
aaccacgtcg cggcactcgc ggaggtagac gcgccgcttc cgccatact ggtggaccgg	20820
aagacgatgc gggctcgtga cgggatgcac cggctcctcg cggctctgct caacggacgg	20880
cagacgatcg aggcgaact gttcgacgga accgcggatg agggattcct gcgcgccgtc	20940
cgggagaacg tgggtcacgg actcccgtg tcgcaggcgg accgccgggc cgccgtgcg	21000
cgcatcatcg tgtcccacc gcactctgtc gacaggcgga tcgccgggc gtccgggctc	21060

ggggcgaaga ccgtcgggc cgtcgggcg agttcaactg ccgtcgtgcc gcagttgaac	21120
acccggttgg gccaggacgg caggggtccg ccgctgaacg ggggcgaggg gcggcgagg	21180
gccatggcgg tactggccga acaccccgac gcgtccctgc gcgaggtcgc ccgtctgtcc	21240
ggggtgtcgc ccgcgacggt cagcgacgta cgccggcggc tggccgccgg cgagtcgccc	21300
ctgccgtcga gacgggaacc ggccgaaccg cggacggcg ccgactccca ccgcaaccag	21360
agcttcgtgg atcccgctcc ggtgctggag aagctgctgc gcgacccctc tctgcggcac	21420
aaggaggcg gccgccagct gctocagctg ctccgccaga acgcggtcgg cgtgcaggac	21480
ctgatggagc tgtccgacgc cgtgccgtcc cactgcagggt ccctggtgat ccattctcgcg	21540
cagcagtacc gggacgcctg gcagtccttc gcggagaagc tggacgagcc cgccctgcgc	21600
tgtccgggt gacgaacggg cggcacggac ccgttcaccg gacatgaccg gcgccgcgc	21660
gcgttcacgg cgcgccgcc gcactccac ggcaccggga ccaccgccgc gtatccggcg	21720
gaccggggcc cggcggggc ggattcagcg ggcgggggc caggtgccac ccgattccag	21780
ccaccgggag agctccgcc ccgagtcctt gcgcacgacc agttcgacga ggccgcgggt	21840
ctggtcctga ccgtgctoga tgcggacgtc ctcgatgttg acgccaagt cgcgatcga	21900
cgtgaacagt tcggccagg gcgcgggctt gtcggagatg gtcaccgaga cggtcgcgag	21960
ctccgtccgg cgcgtaccgg gtttgcgcac gatcctggcg caccgccgt tcccctccc	22020
caacagctcc tcgagctcct cctgcgcgc gcggcgacc agcgggtcgg cgtcggagac	22080
ggcgcgcagc gcgccgacgg cccggcccag gccggcgcg agggagtcca gaacgtccgc	22140
cacggccgtg gcgttgaac gcaggatgtc ccccagagc cgggcgtcac cggccgcgat	22200
cgggtgacg tcggcgacgc cctgccccgc cagccggacg ctgtcctccg ccgctgtctc	22260
cagccgcgc gcgagcagg aggagagcc atggggcgcg tgcgagacga gggccaccgc	22320
gtggtcgtgc acaccggcgt ccatgaccac cggcatgcc tgcacaacg acaccatctc	22380
caggcggtg ttcagcacgt cctgcccggt cagctccgac ggggtgagca cccagggcg	22440

cccctcgaag aggtccgccc gggcggcgag cggcccggaa cgctcgggtgc cggccagcgg	22500
atggcttcct atgtagctgg ccgggtcggc ccgcatcgcg cgcacgtcgt cgtgcgggac	22560
cttcttgacg ctggcgacat cgaggtaggc tcgggccagc ccgctctcct gtgcgcgcgc	22620
gagcacgcgt ccgacctgtg ccggggggcac ggccagcacc gccaggtcga cctgacggtc	22680
cggctcttcc agggatcccg cggccatcgc ctccgccgtc ctggcggcgt tccggtcgac	22740
gtcctccagg tgcacgccga ccccgcgcg ggctcagcgc agagcgacgg acgtgccgat	22800
ggccccggtg ccgatgactg tgggtgtcct caacgcgcgc cccaggtgc ggtgatccga	22860
aatcggctcg gacaagtgc gtgcccgga cgggaaaagg gaattccat ggcgcggtgc	22920
gcgcgaatt taacgttcg gcgcgcgtgt tcaactcgg cgctgcagcg gtogaacaca	22980
gtagcggtag accggaccat tgaggcatcg tgctcagttg gcgacaccgg gtccgataaa	23040
cgccggaatc cgaggagttg acgttgtagt cagcgctgag acacgacgac ctgcatccga	23100
tagaagaagt ggaaataagt tcgctctcca ccgacggctc cccgcggatc gacggggaga	23160
gtcccgagca cgtggaaatg ctggccggcg ccgacaccgc gttccaccg atcatggtgc	23220
accgccgcac cgggcgggtc atcgacggca tgcaccggct gcgcgccgcg atgctgacgg	23280
gocgtacgac gatcgcggtg aggttcttcg acggcaccga ggaggacgcc ttogtcctcg	23340
ccgtgaagtc gaacatcgcg cacggactgc cgctgtccgc cccgaccgc cggcggggcg	23400
ccgggcgcat catggccacc catccccgtt ggtcggaccg gatgatcgcc tcggtggtcg	23460
gcacctccgc caggacggtc gccgagatcc gccgcgacgc cggcgccgcc ggggcggggg	23520
agcccacccg catcgccgg gacggcaggg tacggcccgt cgacgtgagc gagggccgca	23580
gactggccca cgacatgac gtccgcgacc cgggcctgtc gctgcgccag gtcgccgcg	23640
ccgcgggat ctgccggag accgtcaggg acgtcagaca ccggatgctc cgcggtgagg	23700
acccggtgcc cgcgccgcgg ccgcggacc tggtggagcg cggcgccgac cgcggggcgg	23760
agccggccgg gaaggccgcc gcgccgtgc ggacggagcc gcccccgc gtcgtgatga	23820
agcggctgag ggccgatcc gcgctgcgtc tcaacgagaa cggacgcgac ctgctgcggc	23880

ttctggatat ccacacggtc cggctggagg actggaaccg cattatcgaa agcgtgccgc	23940
cgcaccgtct ggagacggtg ggcagcttg cagctcctg cgccgacaaa tggctccgaga	24000
tgcgtcacg catcgaaagc aacgcatcac atctggccgg gtgaacgagg aaacacacga	24060
atccttcgag gagccgtcgg agaaagcggg acggcccgtc ggaacaccct tgtggagggg	24120
caatggagat acggtcgatc gatcacgtcg aattgttcgt cgaggacgcc caggacacgg	24180
ccggcaggct gtgcgactcc ttccgcttcg tccgcgtggg ccgcggcgcc gggaccaccg	24240
gactgcgcgg ctgcgagtcc gtccgtctgc gccagaacga catcgccctg ctgtgacca	24300
cggccaccga cgccgaccac cgtgccgccg agtacgtgaa gcagcacggg gacggggtcg	24360
cggatgatcg catcggggtg gacgacgcgc gcgccgccta cgccgaggcc gtgcggcgcg	24420
gagccgtccc ggtgccgcgc cccgaggagt tcggggccgc cggcgccgt gtcgtcttcg	24480
cctcggtggc gggattcggc gacgtggagc accgcttcgt ctccggggag gaccccgag	24540
cgccgttcgc gcccttcacg gaggagaccg gcgccacgg ctccgggggc atgctgaagc	24600
gggtcgacca cttgcggtc tgcgtcccg ccggcgaact cgacgggacc gtccgccgt	24660
accaggaggt gttcggcctc agccagacct tcgaggagcg gatcgctgc ggctcgagg	24720
ccatggactc caagtcgtg cagagcgacc gcggcgcggt gacgttcacc gtcacgagc	24780
cggacaccac ccgcgcaccc ggccagatcg acgcgttcgt ggctccacc ggcggggccg	24840
gtgtgcagca cgtcgcgttc ctactgagg acatcaccac cgcggtgcgc acctgcaccg	24900
ggcggggggt ccgcttcctc accacgccgc cgagctacta cgagatgctg ccggggcggc	24960
tgggcccggt cggcgtaccc gtggaggagc tcagcgcgt caacatcctg gccgaccgcg	25020
accgtccgg gatcatgctg cagatcttca ccgagtcgac gcaccogagg cggaccctgt	25080
tctgggaact gatcgaccgc cgcggcgcg agaccttcgg cagcaacaac atccaggccc	25140
tgtacgaggc cgtggagcgc cagcaggcgg cggaggcggc cgaccaggaa tgaggaagct	25200
ccccgcagac gcgtgtggac ccggaggaca cgccctccgg gtccacacgc gtctcgggg	25260

ccagcgtcgg ctacgccccg aggagccggc cgccgtgcag cgtctccccg tacggaaca	25320
catggccggc cccacacctc acgggcgcgga gcagtgccag atccgtgccc ggacggccgg	25380
agagggcgag ccggccggga ccgtccaga ccggtgagaa ctgcacgccg ctacacctcg	25440
aggcgaccag gcggggcccgt ggccggggcgt cggacggcac ccaccgcgga aggacgagac	25500
tgtgcgccag cgggacgtcg tgcagcgcg ggcccgctc ggagcggcgc tcgacctga	25560
ccgacgcctc ggcggtgagc cggccgtgga cggagagcgc gccgtcgaag cggcccccg	25620
gagcgagccg tgagcccgcc cggccgacgg tcaccggcct ggtctggtgg atggcgccga	25680
actgcttggg catgccctgg acccagccgc gcaccatcgg caccggctgg tcgaccagg	25740
cgaacgggca gcgcgccatc ggccggccct cgaacgcgca cccgaggagg atcaggaact	25800
ccgagaaccg gcagacggcc gggtcggcca gctccgcgc gtcctcggag caccactgcc	25860
aggtggcgaa caccggcgcc gccgcacccg gatccgctcc cgcgtccagg cccggcgga	25920
ggaaacgccg tcggcgctcg gggtcgacac ggtagtcgac catgaggatc tcgccggaga	25980
agtgccacgg cggaggcgtg agcatcgacg cctgccccga aggggacagg ggaaggctgt	26040
agccgatggg cccggcgcc ccggccgcgt ccgatccgt cggatgtgtg tgcccggtgg	26100
tggccgtcat gggttccctc cgatctgccg gtccggcggg ccgccggacc atgcctgggt	26160
cagccgtcga gcggcgcggt cgagcagcgg gggcggattg aagctgtagg ccaggcgcac	26220
gctcggctcc cgtgtgccgc cgaagcgga acccgcggtg acgcgcacac ccgccgctc	26280
cgcacgggag agcagttcgt cctcgccgag cccggtgccg caccggagcc agaggaagaa	26340
ccgccctcc ggacggctga tccgcaccgg gaggtccgc gcctcccgca gcgcgtcgag	26400
gagggcgctc gccgcgccc tcagaccgc ccgcaacatt tccagatgcc ggtcgtagcc	26460
gccgtcggac agcagccctg cgacggcgag cgaggtgatg tggttgagc acccgccgct	26520
gcggaacagc ccgtgcgacg cgatccgttc ggccagtgcc ggctccgtca ccagccagcc	26580
cagccggagc cccggcccca gggtcttga gaagctgcc agccgcacca cggcccggtg	26640
tccggcgagg gccgccagtg gcggcggggc cgggggaccg tccgtcaggc ccagttcgcc	26700

gtaggcgctcg tcctcgacga ccaggacgcc gtgctccgcc gccgcctcca gcagccgcag	26760
ccggcgctcc agcggcatgg tggcgcccgt cggattgtgg tgggtcgggg tgaggtacac	26820
gaacgcggtg cggccggtgc cgccttcgcc gcccgcgcg gtcccggcga gggcgcgccc	26880
gagcgctcc ggcaccatgc ccgacgcgtc gagggcgacc cgcctcaggc gcagcgcgca	26940
gtccccgagg atgcgtgcc cgaggtcgta gccgaggccc tccacgagca ccgtgtcgcc	27000
gggcctcgcg aggttggtcg ccagcaggtg gagcgcttg gacgtgccc ccgtgacgac	27060
cacgtgctcc ggccgcacg gggaccgccc ccgcacggtg gccggggcgg ccagctcggc	27120
gcgcaggggc agggcgccc gatcgtgtcc gtagcccagt gccgccgtc cgtactcctc	27180
cagcgcgctg gcgtaggcgt cccgcaccag cccaccggc agcagcgccg gttcgaggta	27240
gccgggcccc aggtcgagga cggccgcggg ggcgacctc tgcaccacac cgcgaccca	27300
ccgccgcgtg tgcgacaacg ggcgggcccgt gccgtacggc agggctccc cgcgggcagc	27360
ggtcatcagc ggggtgtcag cacatggcgc aacgcccgt cgcactgggc cagcggggcg	27420
gagccccggg cgagcgccag gcgcagcgt cggtcgccg gggcggggtc agcccagtag	27480
aaggcacggc agggcagggc gtacacatgg tgctcgcgca gcgcctcca gacctcggtc	27540
ccggtcagat gcctgatcag caccgcctc aactggccc ggctgtccg gtccggcacc	27600
ccggtggtcg acaggtccgc cagcccggcg cgcaccacc accgctgggc ggcatgaac	27660
tcgtgcagct ccgtcagccc gccggcggcg gcgtcctcgg agaagcggcg gaccatccc	27720
aggatcagcg gggagacgcc cagcaggatg tcggagtaga tcttctccac cggcaggccc	27780
aggttctcgg agtggaaccag catgccgacc ttgaggtcga ggtcggcca gagcttgccc	27840
gtgtcctcga tgacaccca gcgcacatcg ctggcgtcga ggatctcgta gtggtcgtac	27900
tgggcgcggg tgtcgaagcc gcggaaggac gtgtcgagg cgaggatcac gccgtgccgt	27960
gcgcactgcc cggccagccg gcgcagccg tccgccgaca cgaaccggcc cgtcgggttg	28020
ttcggcgtgg tgacgaagac acagcccacg gactcgagca gtcgcgggg caggtcgtcg	28080

gcgtgcagcg gatcctcctc caggggcacc agaccaggc ggttgccgcg caacaggtcg	28140
gcgatgttgt cgaaggtggg gtggaccagc gccacggagt ccgtgaccga cgccagggcg	28200
cgggagagga tctccatggc caccgacgag gagtagcagc tcagcacacg gccgggtgcg	28260
gacgggtagc ggtgctggcc gagggccttg aagaaggccg cgtgggcctc gcgttcgagc	28320
tgctcgacgg ggcgcttctc gccgtcctcg aaaagcagcg ggagatcatt gacgatcttg	28380
ctctggccgg gagtgagcgg ctgccgggca tgcccgtcgg cgatgttgaa ctgcctgttg	28440
agtgcgagga attccagttg ggtgaggttc tccgcgcctg atccggcgtg cgcagcgtgg	28500
gccttgcttt gcagtgttcc ggacacaggt atgcctctct gggatgtgag ggtttccaga	28560
agcggagcgg acgtaaata gaaggccact ctacggcctt cgcctccgg ctgaaatgcc	28620
tcttcttttc ggcaccgtgt tcaactgcgg tggatgcgca gtcgaacgag ccgtctcgcc	28680
cgcctatcg gccggacatc gcgttcgac ggtgacgcgc gtgcggttcc cgtgtccaac	28740
tgacctgagg gcgcagttgg acggggccacc ggcacacggc cggccgatcc ttgtcggacg	28800
ggcccgggca cgcgaaagtg gacgtgcggg atctgtgttc cgcgcgggt gtctcttcgt	28860
aagccgtgaa gtggggcctt gatggaattg tcgctcgatg aattcgcgtc gctcgcccgg	28920
gaacggctgg acccggcgt ctgggatttc atcgaaggcg gcgcgggaga ggaacgcacg	28980
ctcgccgga acaccgggc attcgaccgc gtcccgtgc gcccgctgggt gctgcgcggc	29040
gcgggcagcc cgcacaccgg caccacgac ctcgggcgga cgtgggacgc gccctcgcg	29100
gtgccccgg tggcctacca cacgtcgcg gaccggccg gtgaggtcgc caccgtccgg	29160
ggaacggcgg ccgcccggg actcccgtc gtgctcagca ccttcgcggg ccgcacgttc	29220
gaggacatcg ccgccaggc caccgtccc ctctggctcc aggtgtactg cctgcgggac	29280
cgtccctca cccgaggcct catcgaacgc gccgagaac cgggcttcga ggccctggtc	29340
ctcacggtcg acgcgccga cctcgccgc cggctgcggg acctgcgaa cggcttcgg	29400
ctgcccgcg gcacggtccc cgccaacctc ccggtggacg gattcgcgga cccgcggcg	29460
cactcccgc cgcacttcga cccggcctg gactggctcg tggaggagt gctgcgctcg	29520

gtctccgaac tgccgttgct cgtcaagggg atcctcaccg gcgccgacgc ggtgcgcgcg	29580
gccgaggccg gggtggacgg cgtcatggtc tccaaccacg ggggccgcca gctcgacgga	29640
gtgccggcca ccctcgacgt cctgcccag gtcgccgagg cggtcggcgg acgcctcccc	29700
gtcctcctcg acggcggggt ccgccggggg cgggacatcc tggcggcgct cgcgctcggc	29760
gccgacgcgg ccctcgtcgg ccgccgggtg ctgcacggcc tcgcggccgg cggggccggc	29820
ggggtgaccg gcgtcctctc cgtcctcctg gaggagctga cggacgcgat gtcccttgcg	29880
ggcctgagga ccctcgccga catoggcccc tactcgtcg gccgggctcc tgaccacccc	29940
cgccgaagca ccgtggacgc cgggaagggc gcggggagcg accggcgcac cgcgccggg	30000
ggaggggccc ggctgcgcct cgcggacctg caccgagtg tcgccgacc ggtcatggac	30060
accatgaact tcctcaacga ggtgacactg cgctaccccg aggcggtgtc cttcgcccc	30120
ggacggccct acgcggagtt cttcgagacc gagcaggtct tccgccatct gcgccgtac	30180
ctcgaccacc tggccgagca gggccgttcg cccgcgcagg tgcgcgacgc gctgttcag	30240
tacggtccgt ccgccggtgt gatccgcgag ctgatcgccc actcgtcgc ggtggacgag	30300
ggcatcgacg tgtcgccga gtcgatcgtg gtgacggtcg gctgccagga ggcatgttc	30360
ctgacgctgc gcgcgtcat gtccggcccc gacgacgtgc tgctcgtctc cagcccctgc	30420
tacgtgggga tcaccggggc cggccggctg ctggacgtcg cggtagccgc cgtcgaggag	30480
ggcgaggacg gcctgtcgtg cgacgccctc gaggccgcg tctcggcgga gcgggcgcgc	30540
ggcaggcggc cgcgggcccgt ctacgtggtc ccggaccact cgaaccgctc cggcgcgacc	30600
atgccgctcg aggccggaa gtccctcctg gagctggcgc agcggtcga cgtcctcgtc	30660
ctggaggaca gcccgtaccg gcacgtcagc ccgggcacgc aggtggcgtc cctgaaggcc	30720
ctcgaccgga cacggcgagt gatccacctc ggttcctacg ccaagaccgt cttccccggg	30780
gcacgcctcg ggttcgcggt cgcgcgaccg ccggtgctgg cgcgggacgg cggcacgagt	30840
ctgctggcgg acgaactcgc caagatcaag agcatggtca cgtcaaacac ctgccgctc	30900

agccaggccg cggtagcggg cgcgtgctg gagtcggcg gccgtgtctc ggagctcaac	30960
gcccgaacg ccgcccacta cggggaggcc atgcgcttca ccctgcagtg cctggagcgg	31020
gagttcccgg ccgcgcggcg gaccggctc ggcgccgct ggaacgcgc cagcggcggg	31080
ttcttctca ccctccaggt gccgttccg cgcgacaact ccgcgtggc ccggtccgcg	31140
caggacttcg gggatcatctg gacgcgatg togtacttct atccgcaggg cggcggcctg	31200
cacaccctca ggctctccac cagctacctg acccacgcg acatcgagaa gggcatctcc	31260
cggctggccg ggttcatoga gttcagtg cgggaccgg tggcctgaac cgcgcgacg	31320
acgaagggcc ccggccgcgc cggcggggcc ctctctcgg cgcgacgctc aggacggatg	31380
cggctcctcc cagaacatgc tgtcggacga ggcgacgaga ccgcgctga agagcgggac	31440
gtcctcgcg cgtactctc ccagctgcg ccgggtctgc acctgcgagc cctcctgcat	31500
cgcgggctc accgcgacg accggaacag cggggccatg ttctcctct gccgcccggc	31560
catgtcgtcg acggcgtcg cgaactccg cgcgtctgc ttgagccga ggacgctga	31620
ctcggcgtcg gagaggctga agtcgtgct ggacatgcc gccaccagct ccacgaacga	31680
ctccagctcg cgttggtgc tccgggtgac ctctctcgc gtccagaagt aggagtcctc	31740
gtcgacgtgc atgtcgtaga acgaggtcag gaactcgtag aagacgccgt actcccgcg	31800
gtaccgggcc tcgaactcgt cgaaggcccg ccgctcgtc atccgcccc ccagcacgct	31860
gttaggggag cgggcggcca gcagcgcgt gtaggtggc agatggacc cggaggagaa	31920
gacggggtcg acgaagcac ccgcgtcacc gacgagcacc atccccggc gccagaaggt	31980
cgtgtggtgg tacgagtagt ccttgcgcac ccgcagctgc ccgtactgac cggctcgtgac	32040
ccgggtggcg tccgccaggt actccttgat catcgggcat tcgtcgatga ggccgcgcag	32100
cgcgtctcc gggtcgcct gcaccttggc cgcgtctcc cggcggacga ccgcgccgac	32160
gctggtcagc gtggagctga gcgggatgta ccagaaccag ccgtgccga accccacaca	32220
gaggatgttg ccggcgtagg gcgccggcat ccgcttgccg ttctcgaagt agccgaacag	32280
cgccaggctc ttgaagaagt ccgagtacgt gcgcgagccg ccgaccgct tgtggatgcg	32340

gctggtgttg cccgaggcgt ccaccacgta ccgggccgac acctcgtgct ccgtgccgtc	32400
cgggtcggtg taccgcagcc cgcgggcccgc cccgtcggcg tcgtccacga catcggtgac	32460
cgtgcgggtcc tggcggacca cgacaccctt gcgggccgcg ttgtccagca ggatcttgtc	32520
gaatttgctc cgctccacct gataggcgaa cgaggtcggt ccggagacct tggacgagac	32580
ggagaaggag aaattccacg gcttggggct cgcacccac cggaacgtcc caccgcgctt	32640
gtgcggaaaa ccggcggcgg cgagttcgtc ggtgacaccg agcagatggc agatgccgtg	32700
aatggtcgac ggcaagaagc actcgcctat ctggtacctg gggaaggtct ctttttcgag	32760
cagcagcaca ctgtgcccct gcatggccac cagggtcgag agcgtcgacc ccgaagggcc	32820
gccgcccacg accacgacgt cgaattcctc gtgctgtcct gtactcattc ggctcccgg	32880
cacgcactga tgcggtcatc gcgctggtga ctcttttgc agggttcac aggactcaaa	32940
ggggccaaag ggtggacgac attcctgtaa tggacagccg ggcccggcg cgcgcacga	33000
aaggcggcgc cggggaattc ccggcggcgc ccggaccggc gtcgggtgag acgggggtca	33060
actcaccggc gcccggtacc gccacacccc gaacaccgtg aggattccgc tgagtgccac	33120
atagatcgcg gtctccagcc actggaacgc ccagtagcgg ctgctggggt ggtacaggac	33180
gtcgacgtgc aggtcgtgtt cggcggaggca caccgcggtg tcgccgaacg tgccccccgc	33240
gccggtcttg ggccggctgt cgaggcagcc gttgaactcg ctggaggcga gggtcctgcc	33300
gtccgcggtg cgcagcggac tggctctggc gatccacgcg tccggcgcgt ccgggatccg	33360
caccccgccg atgacggatc gcccgccgat actgccagg ttctgogccg agttgatcgc	33420
ctcggccgtc atcgccagcg tcgtcctgtc cggcggcatc aggttgggcc gcaccacgtt	33480
cgggaagaag aactggaagg cgatgaagac caccagcgtc accgccatcg cgggcagggt	33540
ccgccgcagc aggagcccga cgacggtgcc gaacgtgaag gccagcgcgg cgtagccgat	33600
cggggcgatg ttgcgcgcac cgaacacgaa ggtgtcgaac tgctccttga cgacgtcgtc	33660
gaagggccgg gccgcccagg tgagcagggc cgcggccgca ccggtcacga tcaccgaggc	33720

cgcgccgatg agcaggatct tgctgagcag ccagcgcggc cgggtgacgc tctggttcca	33780
caccagccga tgggtgccgt tctogagttc octggcgatc aggggagcgc cccagaaggt	33840
gccgatgagc gcggggatca gggccaggcc ggtcgccagg aacagcaggg tgttctggaa	33900
ggtgctgcgg aactggctcc tggcctgggc gcagttggcc gattgtcgc agttggcctg	33960
gtagacgtca tgggcgtcac ggatgtcccc gccaggtag agcaggtaga cggcgatcac	34020
ggccagcgcg ccggcgccga acagggcctg gacgcggaac tgccgccagc tgagccacat	34080
catcgggtgg cccccaggc tgcggcctcg gtgcgggcgg cgggaacggc ggccgcccgg	34140
gtcatgtagg cgagcacgag ttctctgagg gtgaccggct cggaccggta gggcagtgcc	34200
tcggtcgcgg cgccggtgcg gacgaccgcg ctgctgtgct tgccgctgtg ctgaccgag	34260
atcacctcga tcccgccggg cggctggtcg aactcgccgc gggccgcgac cagccgggcg	34320
tgccgggcca gcagctcccc ggtgtcgccg gcgacctgca cccgggcgtc gcacagcacg	34380
atgagatagt cgcagacctg ctccacgtca ccgaggaggt gcgaggagag gacggcgctg	34440
gocccgagct ccagcacgaa ctccatcagg ttctgcagga acccccggcg cgccaggggg	34500
tccagggccg ccgccggctc gtcgaagatc agcagctccg gccgcttggc cgccgcgatg	34560
gtcagcgcaa gctgcgcgcg ctggccaccc gagagctgcc cggccttctg cccggcgctg	34620
agccccacct ggctgatgcg ccgctctgcc aggaccgggt ccagcccgg gttcatcttc	34680
gcgccgaact tcaggtgctc cgccacggtg aacgcgccgt acaccggcgt gttctgcgcg	34740
acgaacccca cccgggccag gtgcgacgcg ttggccgccg gacgcgagcc gaggacgctc	34800
agtgagccgg acgtcggttc ggtcagcccg caggccaggt gcaggagggt cgatttgccg	34860
gccccgttcg ggccgaccag tccgatgaca cggccggcgg ggacgctgag gtgcacgtcg	34920
ctcagggcga gcttgccgcg gcggccgtac ttcttcgtca agccctccgc gtgaagcacg	34980
ggaggggagt ctgcgtgtgg catgactcca tcctcgaatt ccgccccgtt cacggcatca	35040
gtccaaagca cggttcccgt tgccggcggc cgtactttcg gccggtcggc cacggccctg	35100
ctggtgggcg aactgggtgc ggtacagctc ggagtagaga ccgcccggg ccagcagctg	35160

gtcgtgggtg ccccgtcctt ggatccgccc gtcgtcgatg acgaggatct ggtcggcgtc	35220
ctggatggtg gacagccggt gcgcgatgac gagcgggtg cggccggtca gggcgggtctt	35280
gagggcccgc tggatggcca gctcggactc ggagtccagg tgcgccgtcg cctcgtccag	35340
gacgacgac ggagcgcact tgagcaggag ccgggcgatg gccagccgct gcttctcacc	35400
gccggacagc cggtagccgc ggtcggcgc gaccgtgtcg agaccgtccg ggagctggga	35460
gatcgtcggc cagatccgcg ccgcctcgca cgcctggacg atctcgggct cggaggcgtc	35520
cgggcgggcg tacagcaggt tggcccggat ggtgtcgtgg aacagggtcg cgtcctgggt	35580
gaccacgccg accgtgttct gcagcgagcc gagggtcagg tcgcggacgt cgtggccgcc	35640
gatccgcacc gttcccagg tggcgtcgta gagccgtggc accagggtggg tgatcgtggt	35700
cttccccgcg ccggacgggc cgaccagcgc cgtgagccgg ccggccgggg cgtggaagct	35760
cacgtcgttg aggaccagcg cgccggggcc ctgctcgctc ttgcgtcgcg gcatcaactc	35820
cagtgagggc agggacactt cctcggcgtt ggggtagcgg aaggcgacct ggtogaactc	35880
gacgggggga gcggtgccgt cgccgttcgc cgaggcgagg gccggcaggg ggcgggcgcc	35940
gggacgctcg gtgatcagcg gcttcaggtc cagcacctcg aagacgcggt cgaagctgac	36000
cagcgcggtc atgacgtcgc tctggatgtt cgtcagctgg ttgacggggc cgtacagcat	36060
cagcagcagg gcgaccatgg ccaccagcgt gccgatctgc agcgagccgt cgatgacgaa	36120
ccagccgccg aagccgtaca ccatcgccgt ggtgacggtg gtgagcaggg tgacgaggat	36180
gaacagcagc cgtgcgtgca cgtccatcga gatcgcgatg tcccggacga ggcccgcctt	36240
cttgagaaac tcggcggact cgtcctccgg acggccgtag agcttgacga gcatcgcgcc	36300
ggagatgttg aaccgctcgg tcatcatcga gccagcttg gcgtcgttct gcatgccggc	36360
gcgggccagc ttctccagcc gctgggcgat gatcttccc gggatgaaga acagcgggat	36420
catgatcagc gccgccacgg tgatcggcca cgagaggtag agcatcgccg cgagcaccag	36480
gaccagcgtc agcagcgtcg acagcgactg cgacagcagc gaggtgaggg cctgttgggc	36540

gcccacgatg tcggtgttga tccggctgac cagcgaccog gtctgggtgc gggatgaagaa	36600
cgccaccggc tgccgctgga tgtgggagaa caccgcggtc cgcaggtcga agatgaggcc	36660
ctggccgacc cttccggaga accacgtctg cgtgtagacc gccacgacgt tcagcagggc	36720
cagtccggcg acgagcccgg cgaggccgaa cacgacggac gtcttcccgg ggatgatgcc	36780
gtcatcgatg atcattttga gggtcagcgg gatcgacacc gtgatcaggg agtcgacgat	36840
cgtcgccacc atgaccatcg ccatggcccg gcggtagcgc atggcgtagg gaatgatccg	36900
cttgaacgtg ccggacctga ccggctgcgg gtccaccagt ccttcgacct gcagtcgat	36960
cgtgcccata gtcgggtcgt gtcccacggt cacggagact ctctcagtg tgtgtcgcgt	37020
cgctatgtgt cgcgtcgctg tgtggcgcgt cgctgtgtgt cgtcggggcg tcgcgtcagt	37080
gcttgtcgag gaactcgacg atcgccgcgt tgaccgcgtc gggccgctcg aagtaccga	37140
ggtgcccga gtccgggata tccacgagat cgcagtcggg aacggcctcg gcgacctcca	37200
cgcccagggtg cggcgggggtg atgaggtcgt cggcgaaggc cacgacgcga cagggggcgg	37260
ccaccggcg cagcgccggg cggcggctgt ccatgatgtc ggcccaggcg tgccgggcct	37320
gcgcctcccc gccccggac agctcgaaga cgtccagcca ggcggtcacc gcctggtcgt	37380
cgttgagcgt cgcgggcgag aacatccgga acaccgtcga cgcggcgtcg tacgcggccg	37440
gcagccgcac cccgctctcc accagtgccg tctccgcccg cgtctgggcc gcgcgcgg	37500
cgtccgcacg ggcccgggtg gcgatgagca ccgcgcaccg cacgagttcg ggatgccga	37560
tcgccagctc ctgcgcgac atcgccccc gggagggtgcc cacgatccgg cacggcgcca	37620
gatccagggc ctgcgacagg cccttggcgt cggcggtcac gtccagcagc gactacctgc	37680
ccggcggcgc gtcggacggc gggacacccc ggtggtcgaa gacgaccgtg gagtagcccg	37740
ccgtgtgcag cgcggcgctc tgggtcaggg tccaggcatg gccggccgag cccgagccca	37800
tgatcatgag caccggttcg ccccgcccc cagctggta ggcgatgcgg acgccccca	37860
cggtagcgaa gtgcggggcg cgcggcccc cgggtgtggc catgcgcct ctccctctc	37920
gtcgtcgcgg gggccgccc gttgtaccgg ccggccccgg gggcggctca ctatgcacg	37980

cgccacggg gcggggcagt gtgcgcgggg cacgtccatg gacaccccc ggcccgctc	38040
caactgcggt gatgccctca gttggacacc ggccggccgc gtccaagcag gcccggccga	38100
cggttgatcc gctgtgtgga gctgagccat attgggccgc cgtgagccac tgacgccac	38160
caagtccccg cgcttcttcc gcaccggccg cctggcgccg ccgtcgccg gagggaggga	38220
ccacctgtc agcttcgcac cttcagcca cccggctgac acccgagaag atccggtcct	38280
ggctcgtcga ccgggtgcc tactacgcca ggctgccgc cgaggagatc ggcccgacg	38340
tcccgtcgc gcactacgga ctggactcgg tgtacgcctt cgccctgtgc ggagacatcg	38400
aggacggcct cggcctcgtc gtcgagccc tctgtcttg ggacgtcgc accatcaccg	38460
agctcaccga ccatctcgc gaactgacag ccgactgagg gccttcgagg gggaggacga	38520
tgcgtcga aa ggacctggag aggctgacgt ccggtcagct cggcgtctgg tacgcgcagc	38580
agctcgaacc cctgagcccc gtgtacaaca tcgccgagta cgtggagatc cgcggcgacg	38640
tggacgtcgg gcttctggtg tcggcgtgc ggtctgccct cgacgaggcc cagacctacc	38700
ggctccgctt ccggcaggag gacgccggcc ccggacagta cgtcgacgac tcgctggagc	38760
ttcccgcca cgtcgccgac ctccgctccg caggggaccc gcgcgccgc gccgtggagt	38820
ggatgaccgc cgacctggac cgcgccggg accccctcac cggcccgctg gccgccacg	38880
ccgtgttccg gctgggaccc ggccatgtcc tctggtacca gcgtgccac cacctcgtcc	38940
tcgacgggac cagcctctcc gtgttcgcc gtgggatccc cgggtaccga gctcgaattc	39000
gccctatagt gagtcgtatt acaattcact ggccgtcgtt ttacaacgtc gtgactggga	39060
aaaccctggc gttacccaac ttaatcgct tgcagcacat cccctttcg ccagctggcg	39120
taatagcgaa gaggcccgca ccgatcgcc ttccaacag ttgcgcagct gaatggcgaa	39180
tggcgctga tgcggtatit tctccttac catctgtgc gtatttcaca ccgcatatgg	39240
tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagcccc acacccgcca	39300
acacccgctg acgcaaccc cttgggccc c	39331

<210> 40
<211> 40551
<212> ADN
<213> Secuencia Artificial

<220>
<223> fosmid pXYF24 Sint 騁 ico

<400> 40
atcgaatata acttcgtata atgtatgcta tacgaagtta ttagcgatga gctcggactt 60

ccattgttca ttccacggac aaaaacagag aaaggaaacg acagaggcca aaaagctcgc 120

tttcagcacc tgtcgtttcc tttcttttca gagggtatit taaataaaaa cattaagtta 180

tgacgaagaa gaacggaaac gccttaaacc ggaaaatttt cataaatagc gaaaacccgc 240

gaggtcgccg ccccgtaacc tgtcggatca ccggaaagga cccgtaaagt gataatgatt 300

atcatctaca taccacaacg tgcgtggagg ccatcaaacc acgtcaaata atcaattatg 360

acgcaggtat cgtattaatt gatctgcatc aacttaacgt aaaaacaact tcagacaata 420

caaatcagcg aactgaata cggggcaacc tcatgtccga gctcgogagc tcgtcgacag 480

cgacacactt gcatcggatg cagcccggtt aacgtgccgg cacggcctgg gtaaccaggt 540

atittgtcca cataaccgtg cgcaaatgt tgttgataag caggacacag cagcaatcca 600

cagcaggcat acaaccgcac accgaggtta ctccgttcta caggttacga cgacatgtca 660

atacttgccc ttgacaggca ttgatggaat cgtagtctca cgctgatagt ctgatcgaca 720

atacaagtgg gaccgtggtc ccagaccgat aatcagaccg acaacacgag tgggatcgtg 780

gtcccagact aataatcaga ccgacgatac gagtgggacc gtggtccag actaataatc 840

agaccgacga tacgagtggg accgtggttc cagactaata atcagaccga cgatacagat 900

gggaccgtgg tcccagacta ataatcagac cgacgatacg agtgggacca tgggtccaga 960

ctaataatca gaccgacgat acgagtggga ccgtgggtccc agtctgatta tcagaccgac 1020

gatacagatg ggaccgtggt cccagactaa taatcagacc gacgatacga gtgggaccgt 1080

ggtcccagac taataatcag accgacgata cgagtgggac cgtgggtcca gtctgattat 1140

cagaccgacg atacaagtgg aacagtgggc ccagagagaa tattcaggcc agttatgctt	1200
tctggcctgt aacaaaggac attaaagtaa gacagataaa cgtagactaa aacgtggtcg	1260
catcagggtg ctggcttttc aagttcctta agaatggcct caattttctc tatacactca	1320
gttgaacac gagacctgtc caggttaagc accattttat cgcccttata caatactgtc	1380
gctccaggag caaactgatg tcgtgagctt aaactagttc ttgatgcaga tgacgtttta	1440
agcacagaag ttaaaagagt gataacttct tcagcttcaa atatcacccc agcttttttc	1500
tgctcatgaa ggtagatgc ctgctgctta agtaattcct ctttatctgt aaaggctttt	1560
tgaagtgcac cacctgaccg ggcagatagt tcaccggggt gagaaaaaag agcaacaact	1620
gatttaggca atttggcggg gttgatacag cgggtaataa tcttacgtga aatattttcc	1680
gcatcagcca gcgcagaaat atttcagca aattcattct gcaatcggct tgcataacgc	1740
tgaccacgtt cataagcact tgttgggcga taatcgttac ccaatctgga taatgcagcc	1800
atctgctcat catccagctc gccaaaccaga acacgataat cactttcggg aagtgcagca	1860
gctttacgac ggcgactccc atcggcaatt tctatgacac cagatactct tcgaccgaac	1920
gccggtgtct gttgaccagt cagtagaaaa gaagggatga gatcatccag tgcgtcctca	1980
gtaagcagct cctggtcacg ttcattacct gaccataccc gagaggctct ctcaacacta	2040
tcaccccgga gcacttcaag agtaaaactc acatcccgcac cacatacagg caaagtaatg	2100
gcattaccgc gagccattac tcctacgcgc gcaattaacg aatccaccat cggggcagct	2160
ggtgtcgata acgaagtatc ttcaaccggg tgagtattga gcgtatgttt tggaataaca	2220
ggcgcacgct tcattatcta atctcccagc gtggtttaat cagacgatcg aaaatttcat	2280
tgacagacgg ttcccaaata gaaagagcat ttctccaggc accagttgaa gagcgttgat	2340
caatggcctg ttcaaaaaca gttctcatcc ggatctgacc tttaccaact tcatccgttt	2400
cacgtacaac attttttaga accatgcttc ccaggcatc ccgaatttgc tcctccatcc	2460
acggggactg agagccatta ctattgctgt atttgtaag caaaatacgt acatcaggct	2520

cgaacccttt aagatcaacg ttcttgagca gatcacgaag catatcgaaa aactgcagtg	2580
cggaggtgta gtcaaacaac tcagcaggcg tgggaacaat cagcacatca gcagcacata	2640
cgacattaat cgtgccgata cccaggttag gcgcgctgtc aataactatg acatcatagt	2700
catgagcaac agtttcaatg gccagtcgga gcatcagggtg tggatcgggtg ggcagtttac	2760
cttcatcaaa ttgtcccatt aactcagttt caatacgggtg cagagccaga caggaaggaa	2820
taatgtcaag ccccgccag caagtgggct ttattgcata agtgacatcg tccttttccc	2880
caagatagaa aggcaggaga gtgttttctg catgaatatg aagatctggt acccatccgt	2940
gatacattga ggctgttccc tgggggtcgt taccttcac gagcaaaaca cgtagcccct	3000
tcagagccag atcctgagca agatgaacag aaactgaggt ttgtaaacg ccacctttat	3060
gggcagcaac cccgatcacc ggtggaaata cgtcttcagc acgtcgcaat cgcgtacca	3120
acacatcacg catatgatta atttgttcaa ttgtataacc aacacgttgc tcaaccgctc	3180
ctcgaatttc catatccggg tgcggtagtc gccctgcttt ctcggcattct ctgatagcct	3240
gagaagaaac cccaactaaa tccgctgctt cacctattct ccagcgccgg gttattttcc	3300
tgcgttcggg gctgtcatca ttaaactgtg caatggcgat agccttcgtc atttcatgac	3360
cagcgtttat gcactggtta agtgtttcca tgagtttcat tctgaacatc ctttaatcat	3420
tgctttgcgt ttttttatta aatcttgcaa ttactgcaa agcaacaaca aaatcgcaaa	3480
gtcatcaaaa aaccgcaaag ttgtttaaaa taagagcaac actacaaaag gagataagaa	3540
gagcacatac ctcatgact tattatcact agcgtcgcgc gcagccgtgt aaccgagcat	3600
agcgagcgaa ctggcgagga agcaaagaag aactgttctg tcagatagct cttacgctca	3660
gcgaagaag aaatatccac cgtgggaaaa actccaggta gaggtacaca cgcggatagc	3720
caattcagag taataaactg tgataatcaa ccctcatcaa tgatgacgaa ctaacccccg	3780
atatcaggtc acatgacgaa gggaaagaga aggaaatcaa ctgtgacaaa ctgccctcaa	3840
atttggcttc cttaaaaatt acagttcaaa aagtatgaga aaatccatgc aggctgaagg	3900
aaacagcaaa actgtgacaa attaccctca gtaggtcaga acaaattgtga cgaaccacco	3960

tcaaattctgt gacagataac cctcagacta tctgtcgtc atggaagtga tatcgcgga	4020
ggaaaatacgt atatgagtcg tctggcggcc tttctttttc tcaatgtatg agaggcgcat	4080
tggagttctg ctgttgatct cattaacaca gacctgcagg aagcggcggc ggaagtcagg	4140
catacgctgg taactttgag gcagctggta acgctctatg atccagtcga ttttcagaga	4200
gacgatgcct gagccatccg gcttacgata ctgacacagg gattcgtata aacgcatggc	4260
atacggattg gtgatttctt ttgtttcact aagccgaaac tgcgtaaacc ggttctgtaa	4320
cccgataaag aagggaatga gatatgggtt gatatgtaca ctgtaaagcc ctctggatgg	4380
actgtgcgca cgtttgataa accaaggaaa agattcatag cttttttcat cgccggcatc	4440
ctcttcaggg cgataaaaaa ccacttcctt ccccgcgaaa ctcttcaatg cctgccgtat	4500
atccttactg gcttcgcgag aggtcaatcc gaatatitca gcatatttag caacatggat	4560
ctcgcagata ccgtcatgtt cctgtagggt gccatcagat tttctgatct ggtcaacgaa	4620
cagatacagc atacgttttt gatcccgga gagactatat gccgcctcag tgaggtcgtt	4680
tgactggacg attcgcgggc tatttttacg tttcttgtga ttgataaccg ctgtttccgc	4740
catgacagat ccatgtgaag tgtgacaagt ttttagattg tcacactaaa taaaaaagag	4800
tcaataagca gggataactt tgtgaaaaaa cagcttcttc tgagggaat ttgtcacagg	4860
gttaagggca atttgtcaca gacaggactg tcatttgagg gtgatttgtc aactgaaaag	4920
ggcaatttgt cacaacacct tctctagaac cagcatggat aaaggcctac aaggcgctct	4980
aaaaaagaag atctaaaaac tataaaaaaa ataattataa aaatatcccc gtggataagt	5040
ggataacccc aagggaagtt ttttcaggca tcgtgtgtaa gcagaatata taagtgtgt	5100
tccctgggtc ttctcgtc actogaccgg gagggttcga gaaggggggg cccccctt	5160
cggcgtgcgc ggtcacgcgc acagggcgca gccctggta aaaacaaggt ttataaatat	5220
tggtttaaaa gcaggttaaa agacaggtta gcggtggccg aaaaacgggc ggaaccctt	5280
gcaaatgctg gattttctgc ctgtggacag cccctcaaat gtcaataggt gcgccctca	5340

tctgtcagca ctctgcccct caagtgtcaa ggatcgcgcc cctcatctgt cagtagtcgc	5400
gcccctcaag tgtcaatacc gcagggcact tatccccagg cttgtocaca tcatctgttg	5460
gaaactcgcg taaaatcagg cgttttcgcc gatttgcgag gctggccagc tccacgtcgc	5520
cggccgaaat cgagcctgcc cctcatctgt caacgccgcg ccgggtgagt cggcccctca	5580
agtgtcaacg tccgcccctc atctgtcagt gagggccaag ttttcgcga ggtatccaca	5640
acgccggcgg ccggccgcgg tgtctcgcac acggcttcga cggcgtttct ggcgcgtttg	5700
cagggccata gacggccgcc agcccagcgg cgagggcaac cagccgaggg ctlogccctg	5760
tcgctcgact gcggcgagca ctactggctg taaaaggaca gaccacatca tggttctgtg	5820
ttcattaggt tgttctgtcc attgtgaca taatccgctc cacttcaacg taacaccgca	5880
cgaagatttc tattgttcct gaaggcatat tcaaatcggt ttcgttaccg cttgcaggca	5940
tcatgacaga acactacttc ctataaacgc tacacaggct cctgagatta ataatgcgga	6000
tctctacgat aatgggagat tttcccgact gtttcgttog cttctcagtg gataacagcc	6060
agcttctctg tttaacagac aaaaacagca tatccactca gttccacatt tccatataaa	6120
ggccaaggca tttattctca ggataattgt ttcagcatcg caaccgcctc agactccggc	6180
atcgcaaaat gcacccggtg ccgggcagcc acatccagcg caaaaacctt cgtgtagact	6240
tccgttgaac tgatggactt atgtccctc aggctttgca gaactttcag cgggtataccg	6300
gcatacagca tgtgcatcgc ataggaatgg cggaacgtat gtggtgtgac cggaacagag	6360
aacgtcacac cgtcagcagc agcggcggca accgcctccc caatccaggt cctgaccgtt	6420
ctgtccgtca cttcccagat ccgcgctttc tctgtccttc ctgtgcgacg gttacgccgc	6480
tccatgagct tatcggaat aaatacctgt gacggaagat cacttcgcag aataaataaa	6540
tcctggtgtc cctgttgata ccgggaagcc ctgggccaac ttttggcgaa aatgagacgt	6600
tgatcggcac gtaagaggtt ccaactttca ccataatgaa ataagatcac taccgggcgt	6660
atTTTTTgag ttatcgagat tttcaggagc taaggaagct aaaatggaga aaaaaatcac	6720
tggatatacc accgttgata tatcccaatg gcatcgtaaa gaacattttg aggcattttca	6780

gtcagttgct caatgtacct ataaccagac cgttcagctg gatattacgg cctttttaaa	6840
gaccgtaaag aaaaataagc acaagtttta tccggccttt attcacattc ttgccgcct	6900
gatgaatgct catccggaat ttcgtatggc aatgaaagac ggtgagctgg tgatatggga	6960
tagtgttcac ccttgttaca ccgttttcca tgagcaaact gaaacgtttt catcgctctg	7020
gagtgaatac cacgacgatt tccggcagtt tctacacata tattcgcaag atgtggcgtg	7080
ttacggtgaa aacctggcct atttcocctaa agggtttatt gagaatatgt ttttcgtctc	7140
agccaatccc tgggtgagtt tcaccagttt tgatttaaac gtggccaata tggacaactt	7200
cttcgcccc gttttcacca tgggcaaata ttatacgcaa ggcgacaagg tgctgatgcc	7260
gctggcgatt caggttcacg atgcggtttg tgatggcttc catgtcggca gaatgcttaa	7320
tgaattacaa cagtactgcg atgagtggca gggcgggcg taattttttt aaggcagtta	7380
ttggtgccct taaacgcctg gttgctacgc ctgaataagt gataataagc ggatgaatgg	7440
cagaaattcg atgataagct gtcaaacatg agaattggtc gacggcccgg gcggccgcaa	7500
ggggttcgcg ttggccgatt cattaatgca gctggcacga caggtttccc gactggaaaag	7560
cgggcagtga gcgcaacgca attaatgtga gttagctcac tcattaggca ccccaggctt	7620
tacactttat gcttcggct cgtatgttgt gtggaattgt gagcggataa caatttcaca	7680
caggaaacag ctatgacat gattacgcca agctatttag gtgagactat agaatactca	7740
agcttgcatg cctgcaggtc gactctagag gatccacgg ttctttctcc ggtccaccc	7800
ccggttccgg acgccgtacc ggccgacgat cctgctcggc gtcgccatcg cgatcctcgc	7860
cggcttcacc ccgtgaacg agctcgcggc gctggtgaac atcggcacc tgttcgcctt	7920
cgtgatcgtc gcgatcagcg tgatcatcct ccgcaggacc cggcccgacc tgccccgcgc	7980
cttcgcacg ccctgggtgc ccgtgctgcc gatcgtctcg gtcgccgct ccctgtggct	8040
gatgctgaac ctgccggcg agacctgggt ccgcttcggc atctggatgg cggtcggcgt	8100
cgtcgtctac ttctgtaca gccgcaaaca cagccgtctg gccgaggagc gcggcgggga	8160

acggacgtcg tcctgaggcc ggcgccctccc ggcgctaccc gtcccgccgc ggcgcccagg	8220
cgccgccccgc gggccgcacc gcgcgcgggc cggtagctc agcgcccagg gacgtcacc	8280
ggcggcgcag ctgcgggtcg gccgtgacca ccaggaccgg gcggtcgccc gcctccgcca	8340
ccaggtcgac catgcggtcg tcgccgtgc ccggggccgg gtccaccgg acaccgggga	8400
ccgactccac cccgcgggcc gcccctcgg tcaccaggac gatctccacc gggcccgggt	8460
gaccggcac gccctccgc gccagccgt cgcgcagccg ttccggcg ccgcggcggt	8520
ccgccacca tccgtcgggc accgaccga ccacgttgcc ggctcgacg atcacgagca	8580
ggccggtgtc atccatggcg tcagggtccc acgcggcgcc ccggccgtcc gggtcgcggg	8640
ggtgcgaccg cgcgcggggg aggacgccg ggaacacatg aacgatcgac atatgatgtg	8700
tcgagtttca gcgttccgc gtcggtgac cggcgggccg caggaagggg cacctggcga	8760
tgcggagaag gtcggcgggc cccgtggcg cctccgtcaa ggacggacgt gccgccgggg	8820
agcaccggga ggccgcgtcc ggggccgcc acggtgactg gtcacccgc ggcaaggacg	8880
gccggctgac gctgtacgtc cccaccgac gcggtctgt gcgctggac gagaccgccg	8940
tgggtggccc cggctggagc ggtccgact tcgtcccgt ggccgggctg acgcacctgg	9000
cgggtggtca gggagccaac ggctacgtc acttcctcgg ccgcaggag cgcgaggcg	9060
ccgactccac gccggcggtg gacatcgtgc acgcgatcca gtaccagacc ggactcgcct	9120
tcagcgactg gcggtccctc ggcaaccgc accgggtccc ggaggagccc ggaccgctcg	9180
ccgtgccggt cggggcggtc gcccgggac gcaccgtgca cgtgttcgtg cggggcgcgc	9240
acggagggct gatgctgcgg cgcgaggccc cgaacggcaa gtggaaggcg tgggaggacc	9300
tgggcggcgg cggcgccggc gccagccc cggcgctcgc gtcaccgac gggcggatcg	9360
aggtctgcgt ggccgccgag acgggcgtgc tggcgtggag ccagtccaag cccggcgcg	9420
acttacccgg gccccgggt ttctcgtgc gcccgtgcc gggcacggtc gcggccctgg	9480
agaccggtcc gggccgggcc acgttcttct ggacggacgc cgagagcggc ggtgcggcgg	9540
cctggcgggc gggggcgtgg cccgtcgcgc tggcggtac cccggccgag cggccgtgcg	9600

cgggtgctccg cacgtcgctg gacggctacg actgtgtcgt cctcgcctac cgtgaccagg	9660
acggcacggc cgtgctcggc atgggcggca cggagaacga ggccgccggc ttctggtggt	9720
acgcgctgac cgagtcctgc cagggcgctc cggccctggc cctggacggg cggggccgcg	9780
tggtgatggc gctgatcggc gccgacggca ggcccagggt cggccgccag gaggacggcg	9840
acggcctctc gctcaccggg tgggacgtcc tcgggggtg agcgcgccgt cgccgtgccg	9900
gaggtcagcc gcccgtcgcg ggcgacttct tcgcggactc ggggatctcc gcgtcggagc	9960
ggatcgctt ccacagctgg tcggcctgcg ggtgcgcggc caccaccgg ttogggtcgg	10020
tcttgtcgtc cgccacgggg agcatgacgg tctccatgga cgccgggtcg accccgttca	10080
tgctgcgcgc gaagtccgcc aggttggtca gcgaggccag ctcgagtcg gtggtcaggg	10140
ccgaggtcag ggtgtcggcg atcttgtacg tcttggtggg gctgccgaac aggtcctgct	10200
tcttcacctc gctcagcagg gcgatcatga actgctgctg gaggcctatg cggcccagg	10260
cgctgccgtc gccgacgccg tgccgggtcc ggacgaacgc cagggagtcg gtgccgtcca	10320
gcttgtgcgt cccggcgctc aggtccaggc cgctcgtgct gtccttgatg ggctcgtcga	10380
cggtgaccgt gacgccgccg atcgctcga ccagccctt gaagccggcg aagtcgatct	10440
ccatgtagtg gtccatgcgg acgccggaca tcttctccac ggtcttgacc acgcaggccg	10500
ggccggccgt cgagtacacg gattgaaca tgacgcgctc ggcctgcgga agggcggagc	10560
cgtccgcctt ggtgcactcc gggcgggtga cgagggtgtc gcgcgggatg ctcacggcga	10620
cggcctgccg gcggccctcc gggatgtgca tgaccagcgc ggtgtccgac cgggcgccc	10680
ccaccttgcc ggtgcccagg ccggcgttgt ctccggcacg cgagtccgag ccgaggacca	10740
gcacgttctg tccggaggtc ggcagcttct cggggcggtc ctcgccagg gcctcgtcca	10800
ggtcgacccc gtcgatgttc ccgttcaggt cgctgtagag ccagtagccg gtgcccccg	10860
cggcgagcac gacggccagc agggacagca ggacgatcgc tcccagcgc cggcgcttcg	10920
gccggcggcg tccgccgccg gactcgggct gcgcacgcc ggactccgtg gtggcgctgt	10980

gcgtcatggg tcttgggttc ctctcctcac gggctccgcg cggcggagtg gggcccgggt	11040
ggggggagac cgggccacag atgcgttcac gtcagtggga gaactatagg cagcgggtcca	11100
cggcacatcc tgcacgggtg gaattagcac ggctcagcga tgaatgccac atgaaccgac	11160
cttgaccact ctttaaggctg gaataagatg tgtcgggcct gtgaccgtcg tggagggggg	11220
ctttcacgag gtcggagggtg ctggtggggc ggagcggcgc gaagccgcga agcctcacat	11280
aagggtggg cgtccggtgt ggacgatttg gtgtccggc gggcgaaatg tcttggtagg	11340
agtacgagtt ggtggttcag ccgcgtctga gtattgtcgt gcccttcag gacgtcagg	11400
tgtacctcgc cgagtgtctg gaatcgatcg cgcggcagtc gttccgcgac ttcgaggtca	11460
tcctggtcga cgacggctcc accgacgggt ccgtgcggat cgcggcggac ttctgcgccg	11520
ccgaccgccg tttccggctg gtccgccagc acgcccacgg accgggccac gcgcgcaaca	11580
ccggactgcg gaacacgcac cccgcggcgg agttcctcgc cttcgtggac ggtgacgacg	11640
tcatccccga gtacgcctac gaactcctgg tgcgcacgct cgaggagtcc gagtccgact	11700
togtctcggg caacgtgcag atgatgaact ccaccaagaa gtggcagtca ccgctgcaca	11760
agggcccat gcagaagaac cggcgcggga cgcacatcac gaagttcgac gcgctgatct	11820
acgaccgcac cgtctggaac aaactcttcc ggcgctcctt ctggaaccag aactccatca	11880
ggttccccga aggcgtgctg tacgaggact cgtgggtcaa catgtacgcc cacttcgcgc	11940
ccgccaaggt cgacgtcatc acggacgtcg tctatttctg gcgccgccgg gacggcggag	12000
cggcgccctc catcaccacg cgcactccg aactgtcgaa cctccgggac cgggtcgcgc	12060
ccgtgcagtc ggtgagccgc ttctcggcg accggcgctc gcgtgagtac gcggacagca	12120
agcggaagta cgatctgcc tgccctgaagt ccgacctcct gctccatctg aaggtgctcc	12180
cggacgcgga cgaggagtac cagcacgcct tcatgaagtg ggccaacgag ttctcgcacg	12240
agacggatct caccatcatc gacgagctgc ccgcggactc ccggtcaag tggctcctgg	12300
tgcgcgagga gcggctggcc gaactgctcg aggtcatcga gttcgagcgc cgcggcggtc	12360
ccatgcccggt gcagcggcgt ttccggcgct acctgaacta cccgtacctc ggggaccggg	12420

gggtgggcct cgacaagaag gcctaccggc tggacaagga gctctcgctg cacggctcgc	12480
tgtccggagc cgcctggagc accggctccg acctgctcac cctcaccgga acggcgtacg	12540
tccgcttcat caacgtgcac aagaagcaca tgtcggtgaa ggcgatcgcc ctgcggaaca	12600
agaagcaggg gcgcatgcag atcacgacgg cgaagaccgt ctacgcgccg caggcgaccg	12660
aggacagtaa gcagaatcgt tactgctatg actgggcccg cttcgaggcg cgcctcgaca	12720
ccacccgcct caagcgcaag gccagtgagg tcgagggcac ctgggacgtg gccgccggtg	12780
tcctcagccg gggactgttc cgctaccggg gcctcgaccg gggcggcgcg ggcagcgccg	12840
ccaacccgcc ctaccgctac gtcgacaaga acacccgcat cctcccgtc ttctccagg	12900
gcaaactcaa gctgcgcgtc gagatcgtgc gctgccgat caccaagcac cgtgtcgtcg	12960
gcgaccagct ggagctgcgc ggcgtctacc tcggcccca ggtcccggag tggggcaagc	13020
tcgcgctcac cagcatgagc ggcgcgggac gccacgacgc acgcgtccac ttcaccccg	13080
gcggtgaggg ctggtgcacc ttctccgcca agctccccct gagccgtctg gtgccaagt	13140
cccgcgtcca ggcgggaacc gacgcggacg tccgcagtc ctggggcatg ggcagcaacg	13200
gctggaagac caccttcac gtcgagggcc gcaagtcggc catctatccc gtgatggcgg	13260
aggagacccc ggacgggcac tactccatgc cgtcctccct gcagaccccg gagcgcgacc	13320
gggagatcgt cgtgcaccgc aacggctccg gctatctcgt gctcttcgaa cgagcgaccc	13380
tgccctggc gaccgggtgc gactggcagg aggacggctc gctgtggatc cagggccgtt	13440
acctggccgc ggaccagctg accccggagc agtaccgctc cggccacctc gtggtgcgct	13500
cgcgcgccc cggcgcgga cgctccgtac cgctcacctg ggacgggcac gagttccgct	13560
gogtcctggc ccccgccgcg atgoggaccc tggccgggga catcccgtg gcggccggac	13620
ggtgggactt cttcctgcgc cgccaggacc tgtcggccgt ggcccgag gaccggctcg	13680
aagacctcat ggtgaagatc gagcaggatc tcatcgaggc gttccgcag gactacgaga	13740
gaaacgaacg ccgctacgag acgcaggccg aggcctacga ccggtgtcg ctgctcgtcc	13800

actcggcgat gcccgaccac gcccgcggcc cctaccggca gaagctcctg aggaccaagg	13860
cctaccccgga cgcccgggcg cggcgggtgc gtgacgccgt gctgttcgac gccttcaagg	13920
gcaccagta ctcgacagc ccccgcgccc tgcacgagga actcgtgcgc cgccgcaccg	13980
gcctggaaca cctctgggtg gtgcgcgacg accaggtgca ggtgccgccc acggcgacgc	14040
ccgtccgcat gtggtcgccg gagggttacg aggcctcgc caccagccgc tacgtcgtcg	14100
ccaacaacca cctccggac tggttcaaga agcgggacgg acaggtcgtc gtgcagacct	14160
ggcacggcac gccgctgaag aagatcgcc acgacatga gtccatccac ttgcggacc	14220
agcgctatct ggaacgcgtc gagaaggagg tgcagaactg ggacatgctg gtgtcgcca	14280
acagcttctc cccccgac ctcaagcgc ccttcggctt cccggcgag atggtggaga	14340
gcggtaccc gcgcaacgac atcctgcgc gcccgacac cggggcccgg gagcaggaga	14400
tccgccgag catcgggctg ccggagggca agcgggtggt gctgtacgc cgcacctggc	14460
gcgacgacca gttctacgc cccggcaagt acaagctgga cttccggatc gacctggccg	14520
ccgcgcgtgc gcagctggc cccgaccac tcctcatggt gcgcgccac cccaacgtcg	14580
tggaccgggt gccggcgcc ggcgacgat tcgtcttga cgtgtccgac taccggaca	14640
tggccgacct ctcgatgac accgacgtga tgatcacga ctactoctcc ctgatgttcg	14700
actacgtgaa caccggggcg cccatcctgt tcttcaccta cgacctggac cactaccggg	14760
acacctgcg cgggttctac ttgcacttcg agggcagcgc gccgggccc ctctctaca	14820
cgtccgagga actggtggcg gcgatccgtg acatcgacgc catccaggac ctctacccg	14880
agcgttaccg ctggttcag cgggagttct gcgacctgga cgacggttac gccgcggccc	14940
ggctcgccga ccgatgctg gtgcggggcg gcgacctgc cccgggcag gcgcacgcgc	15000
cggccgtcgg cgcggtcgac acccggcaca ccggaaggcc gatgaccccc ctccagtggg	15060
ggaactcgga gtggttcgcc gggcccgcc cgccggcggg tctcgtcgac gccgtgcccg	15120
cccagcccgc cccggcgtac gacgccgtac cgcagacca ggcggtccg ttggccata	15180
ccccgccgc cggcgaccgc agctacgaag gcgtgatcgc gtgacgccgc cgacccgcc	15240

cggcgcgagg tccccggcgt gccgcgcggt gccggaaggc cccggccccc gatgaccgcc	15300
gtgccgttcg ccacgcgggc gccggccgcc gctgccggtc cgtcgtccgg cggccgcccg	15360
ggcgcgtcct cagccggtgc cgggtgccac ccggatcagc agcgcgcggt ggtcggagac	15420
cccgggtgcg ctcaccggc acccgagcac gggcagcccg gtgaagaggt agtccagctt	15480
gtggtgcgag acgtgcgtcg gccgggcccg ccggagggga cggggcgtcc cgtcgcactc	15540
ccggtgcgtg ccgtaaggct ggtcgggccca gaccgggag agcgggttgc gctctcccgg	15600
cggatccacg ttgaggtcgc gcccgtagac ggtgcgccgc tcgggcaccg cgtccaccag	15660
ggccttcagc tgtccggcgc ggaactcgc gtccggatgc gccagatcgc cgcgcgcgg	15720
ggtcagatgc gcggtgcaca ccgtgaggtc gtgcgccgcg acgaacgcgc agagtattcc	15780
gcgctgcacc ccgaccggc gctggggcgc gggcaccgcg cgcacggacg acagcggata	15840
cgcgcacagc agggcgtagc ccgcggagcc ccggccgggc gcccgcagc gcaccgcggt	15900
gcggcgcccg tcacggccgc gccagggtga ggccctgaac tccgcgtgcc acgacgcccc	15960
gagggaggcg cgcaccgcct cgacgtccgc cgcgcaggtc tcctgcagga acagcaccgc	16020
ggccccggat tccgcggcga ggcgttcgt ccggccgcgc ttggcgtcct cgcgcgccgt	16080
gccctcgcag ttccattccc tgaccccgca catgttcag gtcgccacgg tcagtgtccg	16140
gtccccggtg gcggagcgt ccgctccatt gccgtccgc tcgtgggtca ccacggcggc	16200
gagccccgc agggcggcgg ccgccgtgac ggcggcgagc aggcgccgtc cgcgcggcg	16260
gggcgaccgg gttcggttcc tgagcaccgc gccatcatga ccgatccgc gctgaccagc	16320
caaaaccacc cggcgccagg cgccttcgg caggtgtcct cccctgccc gcgcaccgc	16380
ccgtcgtcca agtacctctg agagtggagt ccgtacatgt ccaaggcacc ctogaacggg	16440
cggcagctgc tcaacggcat cgaagcctcg ggaacgttcc cggaggagta ccggttcacc	16500
cacccaaga gcggcaaccg gcacctcgt gtggtcttcg ccaacttctc ggcacccgag	16560
gactacggct ggtcgaacgg cgtcttcgac aacgtccgtg ccaacatcct gtggatccgt	16620

gaccggttcg acgggatgaa cgcctactac ctgtgccgga acatggactt cggctctggcg	16680
gactcgggtgc agaccctgat cgcgaacgtc accggggcgc tcgggctgac gccggaccaa	16740
gtcacgctct ggggcggctc caagggcggc agcgccgcgc tgtacttcgg cctgcggtac	16800
ggctaccgga acatcgtcgc catcgtcccg cagttcctca tcggcgacgc cctggagaag	16860
cggcaccgga aggtctccgc gtacatgctc ggccaagggg cgcaggcgca caacgcgcgg	16920
atcctggacg cgctcctgcc cgacctgggt cgcgccaagg ccaaccggg cgccaacatc	16980
tacgtgctct cctccccgca ggacgagcat tacgccgtgc aggtcgagcc gttcctcggc	17040
atgttcacg gctacgagaa cttcaatttc ctgtacagcg agtcgccgac catcacgggg	17100
cacgccacgg cgaccggcg gaacgtcccg gcgctggtcg gcctgctcaa cctgctcgcc	17160
gacggctacg ccccccggct gggcttcacc cgccacgcc cggaggactt cgaccacgac	17220
cggtcggaca tcaacgccta cctcgcctcg acctccaagg tccaggcgcg cgacgcgttc	17280
gcgccgccgg tggtagccac cccgggcttc aacagcgagg tcccgcgcac cggaccgtgg	17340
ttcaccggga cggcccacgg agcggcgcgg gtgagcatgt ggcgcaacgg caagtctgtg	17400
gcgtcgcccc aggtcgcggc cgacggcacc tggctcctggc agccgaccgg gccgtgggag	17460
gocgggaagc acatcgtcaa gatcttcgcg gtggaccggc cgggcttcca ctccgccgg	17520
gtcgagatcc ccttcaccgt ggtcgaccgg gatcccgcc ctgccccgcc ggtcgtctcc	17580
gcaccggtgt cggggcagca gaccggagcg gcggtcgggt tccacggcag cgcgccggga	17640
gcgtcacagg tcggcttcg ggagaacggc gtgctcctcg gcgcggtggc cgtcgcgcc	17700
gacggcacct ggggctggga ccccgccgg ccctggcccg aggggcagca cctggtcgag	17760
atcgtcgcgg tcgacgcgta cggcatggag tccgcgcccg ccgcccgg cttcaccgtg	17820
ctcgccacg cggcgccgc cggacacttc acgcccggt actgaccgac ggcccaggac	17880
gacgacacac cacgagtggg aagcagacat gccgaaagaa gcgccgacaa cacgcgagct	17940
gatcaccggg atcgacacct ccggcgcgta tccgtcgag taccggttca cgcacgcaa	18000
ggggggcaac cggcacctcg tcgtcgtgtt cgccaacttc gcggtcaagg acgactacgg	18060

ctggtccaac ggcgctcctca acccgggtgcg ggccaacatc ctgtggatcc gtgaccggtt	18120
ccgcgacatg aacagctact acctgtgcga ggggatggac ttctccctgg agcagtccgt	18180
gatcgggctc atctccaagg tgatgaacgc cctggaactc acccggagc aggtcacgat	18240
gtggggcggc tcgaaggcg gcagcgccgc gctctacttc ggcatgcgct acggcttcgg	18300
caacatcgtc tccatcgtgc cgcagttcct cgtcggcacc tatgtgaagc gggcgcacc	18360
caaggttgcc cggttcatgc tgggcgagc ggtgccggag gagaacgtcc gcgcggtcga	18420
cgcgctcatc ccggacctgg cccgttcggg cgtcgcccg cactccaaca tctatctgct	18480
ctcctcgccg caggacgagc agtaccagga gcaggtcgag ctttctctg gactgttcca	18540
ggggtacgac aacttcaact tcgtgttcag cgagtcccc cacatcacc gtcactcgga	18600
cgtcaccgg cgcaacgtcc ctttctgat gggcctcgtg aacatgctc ccgacgggat	18660
gtccccggc ctgggcctgg tgcgcaacgg gtacgaggag ccggaccgc acaggtccgc	18720
catcgagggc ttcttgggc ccacttcggc ggagcggccc agcgccatcc cgatgccgt	18780
ggtgacgcat ccgcttcgc acatggaact gccacggac ggcgtgtact tcacaggac	18840
ggccccggc gcggtgcgg tgagcctgtg ggagcacggc aagtctctg gtgcgctc	18900
ggtggcgccg gacggcacct ggtcctgga gcgggacaag ccgtggagca agggcgacca	18960
tctggtcaag gccgtcgct gggacgcga gaagcgccgc accaaggga ccgtggtccc	19020
gttcaccacg gtcgccggc cgaacgccgc cgcgccggg gcaccggcg ccgcgccct	19080
ggcgccggg cagccgtgc cggcgccgac ggtccacac ccggggcgct acgagcagat	19140
caccggcacg gccgtgcgt tcagcggctt cggccgggc gccgccagg tgggattcag	19200
ggcgggggg accctccttg gcacgagccg ggtcgcggc gacggaacgt gggcctggga	19260
ctccggctgg ccctggcagg cgggcatgca caccgtggag gtgttcgcc tggacgccgc	19320
gggatccgag tcgccgtgg cgcgggtgcc cttcgacgtc atgcacgcca cggcgggcgc	19380
ctcgccgttc gcctactgac cggtcgccac gcgcggaagg gtcctggag atcgccggc	19440

atctccagga gcccttccgc ggtctacggg gactaggcgg gcacgctcgc cgtgcccggc	19500
tocaggaacc gcttcccgtt caccgcgtcg gagacgccct cgcggtccag gtacggcgtg	19560
atgccgccc a ggtggaaggg ccagccggcg ccggtgatca ggcacaggtc gatgtcctgg	19620
gcctcggcga cgacgccctc gtcgagcatg agcccgatct cctggggccac cgcgtccagg	19680
acgcgggcgc ggacctgctc ctcggtgagg acggtgtcgc cctgcttcag gagcgcggcg	19740
acctccgggt ccagctcggg cttgccgtg tcgtagaggt agaagccacg cttgccggcc	19800
tcgacgaccg ccctgagggt cggggagacc gtgaagcgt ccgggaacgc cctgttgagg	19860
gtctcggaca cgtgcagacc gatcgcgggg ccgaccagct ccagcagcac cagcggggac	19920
atcggcaggc cgagcggctc gacggccttc tccgcgacct cgaccggggt gccctcgtcg	19980
atgacgttct ggatctcgcc catgaagcgg gtcaggatgc ggttcacgac gaacgccggg	20040
gcgtccttga ccaggaccgc ggtcttcttc agcttcttgg cgacaccgaa cgccgtggcc	20100
agcgcgcgt cgtcggtcgc ctgcgcggg acgatctcca ggagcggcag gatcgcgacc	20160
gggttgaaga agtgaagcc gacgaccgc tcggggtgct tcagcttcga cgocatctcg	20220
gagacggaga gcgaggaggt gttggtggcg aggatcgcgt gcgccggggc gaccgcctcg	20280
acctccgcga acacctgctg cttgacgccc atctcctcga agacggcctc gatcacgaag	20340
tcggcgtccg cgaagccctc ggccttgtcc agcacaccgg tgaccagggc cttgaggcgg	20400
ttggccttgt cctggttgat ccggcccttg ccgagcagct tgtcgatctc ggcgtagacg	20460
tagccacac ccttgtcgat gcgcgcctgg tcgatgtcgg tcagcacgac cggcacctcg	20520
aggcggcgca ggaacagcag cgcgagctgg gaggccatca gaccggcgcc gaccacgcc	20580
accttggtga ccgggcgggc cagcgacttg tccggggcgc cggccggccg cttgccgcgc	20640
ttctgcacca ggttgaacgc gtagatgccg gagcgcagtt caccgccat gatcaggtcg	20700
gcgagcgct ggtcctcggc gtcgtagccc tgctggaggt cgccgttctt ggcgcgggcg	20760
atgatgtcca gggcgcggtta ggcgccggg gcggcgccgt gcaccttgga gtcggcgatg	20820
aagcgccct tggcgacggc ctggtcccag gcctcgccgc ggtcgatcac cggcgctcg	20880

atccg gatct cgtccttgag gacggccgcg gtccagatca gcgactgctc caggaagtcc	20940
gcgccctcga agatcgcgtc cgcgatgccg agttcgaaga cctgcgcgcc cttgagctgc	21000
ttgttctggt tgaggctgtt ctcgatgatg accgagacgg ccttctcggc gccgatcagg	21060
ttcggcagca gcgtgcagcc gcccagccg gggacgagac cgaggaagac ctoggggagc	21120
gagaacgccg ggaggggcgc cgacaccgtg cggtaggtgc agtgcagacc gacctcgacg	21180
ccaccgcccc tgcgcgcgcc gttgtagtac gcgaaggtcg gcacggccag cgtcgacagc	21240
cgcttgaaga cgtcgtggcc gcccttgccg atggccagcg cgtcctcgtg ccgcttcagc	21300
agctcgacgc ccttgaggtc ggccgcgacg gcgaagatga acggcttgcc ggtgacgccg	21360
acgccgacga tctcgccgtc cgcggcctcc ttctcgacct ggtcgatcgc ggctcgatg	21420
ttgccagcg actgcgggcc gagcgtggtc ggcttggtgt ggtcgtggcc gttgtccagg	21480
gtgatgagcg cgaagcgccc ggcccccagg gggaggtcga agtggcgcac gtgcgcgctg	21540
gtgacgacct cgccggggaa cagctcggcc gcacccttca gaagctctgc ggtggtgctc	21600
acttgtcccc ctgcaagtgc gggttctccc agatgaccgt cgcgcccatg ccgaagccga	21660
cgcacatggt ggtgaggccg taacggacct gcggctgctc ctgcaactgg cgggccagct	21720
gcgtcatcag acggacgccg gaggaggcca gcgggtggcc gaacgcgatg gcgccgccgt	21780
actggttgac gcgcgcgtcg tcgtcggcga tgccgtagtg ctccaggaag gccagcacct	21840
ggacggcgaa ggcctcgttg atctcgaaca gaccgatgtc ggagatggac agccccgcct	21900
gggcgagggc cttctcgtg gccgggatcg ggccgtagcc catgaacctg ggctccacgc	21960
cggcgaagga gtaggagacc aggcgcacct tgaccgggag gccgttctcg cgggcgaagt	22020
cctcgctcgc gatgaccgag gcgggtggcg cgtcgttcag accggccgcg ttgccggcgg	22080
tgaccgggcc gtggacgcgg aaggggtctt tcaggccggc caggttctcc agggtggtgc	22140
ccgggcgcat cggctcgtcg gcggtgacca gggccagcc ggtctcaccg gcctcctcgt	22200
tggtgcggcg caccgagacc gggaccaggt cggcctggat cttgccgttg gcgtaggcct	22260

tggcggcctt ctcttgggag cgcacggcgt actcgtcggc gcgctgcttg gtgatcgagg	22320
ggtagcggtc gtgcaggttc tccgcgggtca tgcccatgaa cagggcggac tcgtcgacca	22380
gcttctcggga gacgaagcgc gggttggggt cgacgccctc gcccatcggg tggcggccca	22440
tgtgctcgac gccgcccgcg atggcgacgt cgtacgcgcc gaaggcgacg gagccggcga	22500
ccgtggtgac ggcggtcagg gcgccggcgc acatgcggtc gatggagtag cccgggaccg	22560
aggtgggcag gcccgcgagg atgccggccg tgcggccgat ggtcaggccc tggtcgccga	22620
tctgctggt cgcggcgacg gcgacctcgt cgatcttctt cgggtcgaga ccggggttgc	22680
ggcgcagcag ctcccggtac gccttcacga ccaggtcgtc ggcgcggtc tcgtggtaga	22740
tgcccttcgg gcccgcttg ccgaacgggg tacggacgcc gtcgacgaag acgacgtccc	22800
tgacggtacg aggcacgatg gctctctcc cagggtcggg gacgtgagc gcttgcttac	22860
gccatgctac ttatgagtaa cgtgactgcc cagtcccggc ccccgagcg gcgaacatca	22920
cacgtacggc ggcgcccgcc aaacgccgga ggggctggaa tcagccccct ccggcgtttg	22980
aggagcggga acccctcacg gtggagccgt ggagcctcgc ggggtcagga ccggggtcgg	23040
taccgggtgc gaccccggt gctcctcccc ggtgatgacg ccgaacagcg tgcgcgccac	23100
ctccgccccg aaccctgca catcgtgact catcgcggac agcgtcggat gcgtcagccg	23160
gcacagctgc gagtcgtccc acgccagcag cgacacgtcg tccggcaccg gcagcccat	23220
ctccgcggcc accgacagcc ccgccaccgc catgatgtcg ttgtcgtaca cgatcgccgt	23280
gggccgttcc cccggcgagg ccgccagcag cgaacgcgtc gcccgcgccc ccggtcccc	23340
cgagaagtcc gtggccgtct gccacgccg cgcgggcggc tccagcgccc ggaccgcctc	23400
gtogaacgcc gccgtcgga tcgaggtgtg cccgagcgcc gccgcgcgc ccaccgggc	23460
gatccgccgg tccccgagc ccgccagata ccgcacggcc tccgtcacgg ccgtggcgtc	23520
gtccgtccac acggaggtga gccgcgccgt cagcgccggg tccccgacgg ccaccgccgg	23580
cagcccgagc cgctccgcca ccgccggacg ggggtcgccg gcccggaagt ccaccaggat	23640
cgagccggcg atctgccgcc cgcgccacca cgactccatc agcccgacct cctcctccgg	23700

gctgcgacc agccgcagca gcagcgagca ggaccgctcc accaggacgc tctccacccc	23760
ggagatgaac tccatgtaga acggctccag gccgagcagc cgggcggggcc ggcagaccgc	23820
gagaccaccc acgtccaccc gcgacccggc cagcgtgcgc gcggttcggc tcggcgccca	23880
ccccagctcc cgcgcgccc ggaagatgcg gtcccgggtc gcctccgaca gcccgggctt	23940
ccggttgaag gcgagggaca cggcgccctt ggacacgccg gcgcgcgcgg cgacgtccct	24000
gatggtgacg cgaggggtcg gcgttgccgt catcgagtgg gctccacgca gtacagggcg	24060
gaacgggcgg tgtccgggtc cggagtcttc ccccccgca cccgatggt cacctgttcc	24120
ccggggagca gggtcaccag cccccggtcg gcccgcccc cggggtccag ccggtcggcc	24180
tggagcagca ggtcccgtac gagggtaggg gccgtgaccg tgatcccgtc cggcgcgagg	24240
gcgacctga actccggcgg ggggtagggg atctcccgt ccggcgccgg gaagtccac	24300
gccgcaccc cgtcccgctc gccgaccagg aactccccg gcccggtccg cagcagttcg	24360
accgggacct cgaccacggc caccgtccgc cccccggct ccagcgccgg ggccgcctcc	24420
gcgatcgggg cgcgctcgac ggacatccgg cgcagccgca gcgttcccc ccagtcctcc	24480
gcggactggt tgaccgcgc caccaccaga ccgtcacgt ccgcgcgcac ggtcagcagc	24540
ccgtccgct acagccggcg cagctcgtgg tagagcggct tctccgccc gtcccgtcg	24600
atcgcgccc acgacgtcac cggccagcag tcgttagct gccagaccac cgtgcccgcg	24660
cacaccggcc agtgcgagcg ccagtgtcg acaccggcg ccaccgcac gccttggtg	24720
acctgcgtca gatagtcca gcggtcgaag tcgccctccg gcacggcgaa gtggcgggcg	24780
aggccgcgt ccagcttgc gttgccgtcc tccgccttct ggtggtgcag catgccggg	24840
gagtccggcg cggggtcctc cccgggcagc gcccgccgca gcgtggcgtg cgogggaggc	24900
gcctgccagc cgaactcggc cacgaagcgc gggacgtcg gccggtagtc ggcgtagtcg	24960
gcgcggttcc acacctcca ggagtgtgg gtgccgtcg ccggatcgtt ggggtgtgc	25020
cgccaggaac cggaccagg actgcccgc gtgtacggc gcgtcgggtc cagctccgcg	25080

accacccgcg gcaggacgcc gaggtagtag ccctcgcccc aggagtcccc ggcgagcccc	25140
tgctcccagt cccagtcccg gaacccccac aggttctcgt tgttgccgtt ccacagcacc	25200
agggaggggt gcggcatcag ccgtacgacg ttctcccggg cctccgcctc cacctccccg	25260
cgcagcgggt gctcctcggg gtaggcggcg cacgcgaacg ggaagtcttg ccagaccagc	25320
agccccaaact cgtcgcaggc gtcgtagaag tcctcgtcct cgtagatccc gccgccccag	25380
accgggacca ggtccacccc cgcgccggcc gcctgctcca gccggtgccg gtagcgctcc	25440
cgggtgatcc gggacgggaa cacgtcgtcc gggatccagt tgacgccccg cgcgaacagc	25500
cgctcacgtg tgacgaccag ggtgaacccg gtgccgtgcg cgtcggccga ggtgtccagc	25560
tcaaccgtcc ggaaccgggt cctgcgccgc caggcgtcca gcgcctcgtc accgtgggac	25620
aacgtcagct cgacgtcgta cagcggctgt tcgccgtatc cgcgcggcca ccacaggcgg	25680
acgtccggca cccggagccg cacggtcccg gccgtcccat cgacccgcgc ccgggcgcgc	25740
acgcccccg cgctcgctc cagggtgagc ggtgcctcga cccgggagcg ctccacgtcg	25800
accgccagct cgatctgccc caccocgtcc tcgacggtga ccagcgggcg caccggggcg	25860
atccgcgccg tcgaccagcg ctccagccgc accggccgcc agatcccggc cgtcaccagc	25920
gtcggccccc agtcccagcc gaacgagcag gccatcttcc gcagggtactg gtacggctcg	25980
gcgtacgtc cggggcgctc gccagcctg ccgcgcaccg cctccgcctc ggcgtagcg	26040
gaggcgaacc gcaccgtgag ccggccgctc agtcccgtca cgtcgaagcg gtacgagcgg	26100
tgcatgttcc gcgtccggcc cagtggccgg ccgtcgagca ggatctcggc gacggtgtcg	26160
agaccgtcga agacgaggtc cgtctgctcg tgcgggcccg tccggcggt cagctccgtc	26220
tcgtacgtcc actccgcgg gccacccag gccacctcgg tctcgttgcg gccgaggaac	26280
ggatcgggga tcagcccggc gccagcaga tcggtgtgca cacaccccg caccgaggcg	26340
gggagggcgt ccccggtgcc gtccgggtgt cgcaggatcc atccctcggg gagcgggtgtg	26400
acctgacgca tgcacactcc ctaaaccggt tgagccttct ctgaagagtg gtctggcatc	26460
gttggcgcga ttgcgacttt accggttcag ttcagggtg ccagagtgcc gaatcagcca	26520

tcccactcgt gctcgtccgt cccgtgaacg gagccgtgat gcatctgaac cgccgtacga	26580
cactcaccgg atcgctcgcc ctgctcgccc tcctggcctc cgctgcacg ggcacggggg	26640
gttcctcgaa gggcgcggaac gccaaaggctc ccgacgaccc gtcaaaggtc aagggggtccc	26700
tcacggtcct caccacccgg accgatctgg tgcaggacgg gacgatgaag aagtacgccg	26760
ccgagttcaa cgagacctat cccgggggtga aggtggagtt cgacggcctc accgactacg	26820
agggcgaggt caagatccgt atgaacacgg agaactacgg cgacgtcctc atgatcccgg	26880
cggctcgtcga gaagaaggac taccogaagt tcttcgcctc cctgggcacc aaggccgaac	26940
gcgccgccaa gtaccggttc accgactact ccaccgtcga cggcaaggtc tacgggcaga	27000
gccccgtcgg cgtcgtcccc gggttcatct acaacaagcg ggtgtggagc gaggccggcg	27060
tcaccgactg gccaccacc cccgccgagt tcctggacga cctgaaggcg atccggtcga	27120
agaccgacgc ggtgccgtac tacaccaact tcaaggacat gtggccgctg acccagtgga	27180
ccaacgtcaa cggctccgtc ggctgcgacc cgcacgccac cacgaagctc gccgagggcg	27240
accctgtggc cgagggggcc gacctgcgcg tgggcgacac cctgctccac gacatcgtgc	27300
ggggcggact cgccgagaag gaccogacca ccaccaactg ggagggtcc aagcccaagc	27360
tggccaaggc cgagatcgcc accatgtggc tgggctcctg ggccgtcgtg cagatgcggg	27420
acggggcgaa gcaggccggc gccgaccccg ccgacatcgg cttcatgccc ttccccgcac	27480
agcgggacgg cacgttctgc gcggtgacct ccccgacta ccagcaggcg gtcaacgtca	27540
actccgacaa caaggaggcc gcccgcgctt ggatcgactg gttcaccgac aagtccggct	27600
acgccgaggc caacctcgcc ctatcccccc tgaaggacgc cccgctgccc gccgtcctcg	27660
agccctacga gaaggccggc gtgaagctcc tggacctcga ggacagcaag ggogccgagg	27720
tgaagtccct cgacaaccgc tccgaggtcg gcatctacaa gcccgactac gccaggaac	27780
tcgtcgacct cgcccgcggc gcccgcaagg gcggcctgga cgactacctc ggcggcctcg	27840
gcgagcgtg ggccgaggcg cgcagcgcgc tgggggcctg atgacggaca ccaccgcaa	27900

ggcggcgcg	cggttcccc	cggccgcgcc	cgcggggcgc	ggccgcgcgg	cgcgggcccc	27960
gcgcgcgacc	cggctgtgc	gccgcctcac	cccgtggctg	ttcctggccg	caccgctggc	28020
cctgtccttg	accttcacct	acgcgcccga	tcgccaacat	ggtcgcgtac	agcttcaccg	28080
actgggacgg	cgtgagcccg	gagctgaact	ggacgggcac	cgggaactac	accgaactcc	28140
tcacccgctc	cgagctgttc	gaggtcttct	tcgtcagcgg	ctactacctc	gtgcctccg	28200
cgggtgcagat	cgtgctcgcc	ctctacttcg	ccacggtcct	cagcttcgac	gtccgcttcc	28260
ggaacttctt	caagggcgtg	ctgttcttcc	cgtacctcat	caacggggtg	gccatcggct	28320
tcgtcttcct	ctacttcttc	caggacggcg	gcaccctcga	ctccgtactg	ggcctgctcg	28380
gcgtcgagac	cgaccacgcc	tggctgggca	cgcggttctc	cgcgaacacc	tcgttgccg	28440
gcgtctccgt	ctggcgctac	ctcggactga	acttcgtcct	cttcctcggc	gcgatccagt	28500
ccatcccggg	cgagctgtac	gaggcggccg	agatcgacgg	cgcgaaccgc	tggcagcagt	28560
tcggcacat	catcgcccc	ggcatcagac	ccgtgctgag	cctgagcgtg	atcctctcgg	28620
tctccggctc	gctgtcggtc	ttcgagatcc	cgtacatcat	gaccggcggc	gccaccggca	28680
cggagacctt	cgtgatccag	accgtgaagc	tggcgttcca	gttcaacaag	acgggactcg	28740
cctcggccgc	cgcctcgtc	ctgctgctga	tcgtcctggc	ggtcacctgg	gtgcagcggc	28800
gcacgtccc	cgacgagaag	gtggacctcg	tatgaccgcg	cgtaccgcgg	cacgcgccct	28860
ggtcctgacg	tccctgatcc	tggcgacgct	ggtggtgctg	ctgccgctcg	ccgtggtctt	28920
cctgacctcg	ctgaagtctt	ccgaggagat	ggcgaacggc	agcggagcgc	tgacgccgcc	28980
cgacgaccgc	ctgaacttcg	gcaactacgt	gacggcgttc	cgggacggcc	agatgctgtc	29040
cgcgttcggg	aacacggccg	tcatcctggt	cgtggccgtc	ggcggaacga	tcctgatcgg	29100
ctcgatgacg	gcgtacgcga	tcgaccgctt	ccggttcgcg	ttcaagaagc	tggctcgtggc	29160
gctgttcctg	ctggccgcgc	tggccccggg	ggtgaccacc	caggtggcga	ccttcagat	29220
cgtcaacagc	ttcggcatgt	tcgacagcct	gtgggcgcgc	atcgccctct	acatgggcac	29280
ggacatcgtc	tcgatctacg	tcttcctgca	gttcatccgc	tccatccccg	tctccctgga	29340

cgaggcggcg cgctggacg gcgccaacgc gttcacgcgc taccgcaagg tgatcttccc	29400
gctgctcaag cggcgatcg cgacggtggt gatcgtaaag gggatcaacg tctacaacga	29460
cttctacatc cccttctctt acatgccctc cgaggacctg ggggtcatct cgacgtccct	29520
gttcgccttc aagggccctt tcggcgcgca ctgggagacg atctcggcgg gcgcggtcct	29580
ggtcatcctg cccaccttga tcgtcttctt gttcctccag cgcttcatct acaacgggtt	29640
catgcggggg gcgacgaagt agccagcgcg gccaccagca cgggagtgcg ctggtcgacc	29700
tgcacgcgc gtgccccgtg cgccgtcagc gcccgggcga cggaccgctc gtogggcccc	29760
ggcggctccc agcagacctt gcgcaccgtg tccggggtga tcaggttctc cggcggcatg	29820
ttcagccgtt cggcgagttc ggcgaccccc gcgcggggcg cgcacagccg ggccgcggca	29880
acggggtcct tgtccgccca ggcgcgcggc ggcgaggggc cggtcaccgg ctggccgggc	29940
tgcggcagct gggcctcgct cagcgcttc gcgcggtcga cggccgcctg ccactgctcc	30000
agctggcgcc gccccacccg ctgccgaac ccgttgagcg cggccatggc gtgcaggtt	30060
gcgggcagcg cgagcgcggc ctccacgac gccgcgtcgg aaagcacctt gccgggggag	30120
acgtcacggc gccgggcgat ccggtcgcgg gtctcccaca gctcccgcac caccgccatc	30180
tggcggcgcc ggcgcacctt gtgcatgccg gaggtgcggc gccaggggtc cttgcggggc	30240
tccggggcg gggccgaggc gatcgctcg aactcctgcc gggcccagtc cagcttgccc	30300
tggcgtcca gtccttctc cagggcgtcc cgcagatcga ccagcagttc gacgtcgagg	30360
gcggcgtacc gcagccaggc ctccggcagc ggacgggtgg accagtcgac ggcggagtgg	30420
cccttctcca ggacgaagcc gagcacgttc tcgaccatcg cgccgagccc gacgcggggg	30480
aaccggcaa ggcggccggc cagctcgggt tcgaagaggc gggagggcac catgcctatc	30540
tcgcgcagac agggcaggtc ctgggtggcg gcgtgcagca cccactcgac gccggacagc	30600
gcctcgccga gggcgacag gtccggggcag gccacggggt cgatcagcgc ggtacccgca	30660
ccctcgcggc gcagctggac gaggtaggcg cgctggccgt agcggtagcc ggaggcgcgc	30720

tcggcgcca cggcgacggg tccgctgccg gccgcgaagg cggcgaccgc ctccggcagg	30780
goggcctcgt ccgctatcac gggcggaatg ccctcgcggg gttccagcaa gggggtcggc	30840
gcctccgtaa cagaagatcc gccgtcgtcc ggaggagcgc ctccggtggt gcgcagtga	30900
ctgtctgctg cgggtgctg ggcgtcggtc acctgtcaag ggtatccgtg ccgcgaaggc	30960
gcccgtcgac gtttgtgctc cgtgacgggc gccggtgggt cgtattccgg tcagaagagt	31020
gaaagaacgt gttcgttgg ccgtgggcgg gcggatcggg gacgggcgga tcgggggcgg	31080
gacggaaggg tcagtggatg atgccggtgc gcagggccac ggccaccatg ccggcgcggt	31140
cgcccgtgcc gagcttgcgg gcgatccggg ccaggtggct cttgacggtc agtgcggaca	31200
ggcccatcga gacgccgatc gccctgttgc actggccttc gccaccagc gcgagcacct	31260
ccacctcgc gccggacagc tcgcggtagc cggccgggtg gtcgggggca cccggggggc	31320
ggcgggtgcag gcgcgcggcg gcggcgccga tgggagcggc acccggccgg gtggggagcc	31380
cgaggttggg gcgggtgccg gtgacgacgt agcccttgac tccgccgcg agggcgttgc	31440
gcacggcgcc gatgtcgtcc gccgcggaga gggcgaggcc gttgggcccag ccgcggcgcc	31500
gggtctcgga gaggagggtg aggcgggaac catccggcag atggacttcg gcgacgcaga	31560
tgtcggggg gttccgatg cggggacgag cctccgcgac ggacgaggcc tcgatgacgt	31620
cgcgtacgcc gagcggccac aggtggcggg tgacggtcga acggacgcgg gggtcggcca	31680
cgaccaccat ggcggtcggc ttgttcgggc ggtaggcgac caggcttgcg ggctgctcga	31740
ggagaacgga caccaggcct cctggggtgc gggacgggcc ggctcgtggg ggtgaaggcg	31800
ggacgaaccg tgctttcaag gtcacagtgc tcttcggcag caaacctggt gtctttaac	31860
gaatgatcac gaagtgatga gtaacaatcc gggcaattcg gacgcacgat cgatcattcg	31920
aagatcgaac ggtttcggtc tgcgtcgcaa cgcttcgaa agtggccgta tcgacaaaga	31980
gagatgcagg aggcgggtcg tcgggacccc gcagcgggag gctcagcgcg actgcggccc	32040
cctccgctgc ggcagcgtca ccacggacgc gtccccgga gcggccggcg gcagccccgc	32100
gacctgcgcc agcagatcgg accacgcgac cagatgggcg gccgtgtccg gaaccccgcc	32160

cagaccctca cgcggcgtcc acgaggcacg gatctcgatc tgggaggcgg cgggccgcgc	32220
ggacagcccc ccgaagtagt gcgaactcgc ccgcgtcacc gtgccgctcg gctcgccgta	32280
cgacaggccg cgcgccgcca gcgcgccggt cagccaggac cagcacacgt ccggcagcag	32340
cggatccgcc gccatctccg gctccagctc ggccgcgacc agcgtcacca gacggaaggt	32400
ccccgccag gcgctgtgtc cggccgggtc gcacagcagc accagccggc cgtcggccag	32460
atcctcctcg ccgtcgacga ccgccgcctc cagcgcgtgc gcgtacgggg cgagccgttt	32520
cggcgcgggc accgtctcca cctogatctg cggccgcagc cgggcgctct gcagcgctc	32580
gacagcggcc cggaaggcg gcggaggcgc acctccgtgc ccgggatccc cccgccctc	32640
cttcggttcg tccattccgc cagcgcctc cgacagtcgt ccctgagccg cagccatgcc	32700
gggaagatta agcggaacgg gccccggcg cagggaggga caccgcgcc gcccggcgct	32760
gtccgatcc tgcaccgcgg ccccgccgc cggacgccc tggggtccgg ggcggcggac	32820
gggtcgtcgc agactggccg gtgtgagtgc caacacgagc ccgaagggcc agacgcctac	32880
cgcgaccccc gaccccgta agaacgacgc cgtccgggaa tcagccttcc tcaaggcgtg	32940
ccggcgcgag ccggtgcgc acacgccggt gtggttcatt cggcaggccg ggcgctcact	33000
gccggagtac cgcaaggcgc gcgaggcat cgggatgctc gactcctgca tgcggcccga	33060
gctggtcacc gagatcacc tccagccggt gcgccgccac cactcgacg cggcgatcta	33120
cttcagcgac atcgtcgtcc cgtcaaggc catcggcatc gacctcgaca tcaagcccg	33180
catcggcccc gtcgtcgagc agccggtgc caccgcgcc gacctcgccc ggcgtgcgga	33240
cctgaccccc gaggacgtct cctacgtcac cgaggccatc ggcatgctga ccgtgagct	33300
cgggtccacc ccgtgatcg gtttcgggg cgcgccgttc acccttgca gttacctcgt	33360
cgagggcggc ccgtcccgta cgtacgagaa cgccaaggcg atgatgtac gcgacccga	33420
gctctgggcc gacctgctc accgcctcgc cgacatcac gcggccttcc tcgacgtcca	33480
gatccgggcc ggccctcgg ccgtgcagct cttcgactcc tggccggcg cgtcgcgcc	33540

ctccgactac cggcggttcgg tgctgcccgc ctccggcgaag gtgttcgcgc cggtagccgg	33600
ccacggcgtc ccgcgcatcc acttcggcgt cggcaccggc gagctgctgg ggctcatggg	33660
cgaggccggc gcggacatcg tcggcgctga ctggcgcgtc ccgatggacg aggcgcgccg	33720
ccgcgtcggc cccggcaagg cgctccaggg caacctggac ccgaccgtgc tgttcgccgg	33780
ccgggaggcc gtcgagacga aggcgcgcga ggtcctggac accgccgcgg gcctggaggg	33840
ccacatcttc aacctcggtc acggagtgat gccctccacc gaccgggacg ccctcaccg	33900
tctcgtggag tacgtccaca cgcagacggc gcgctgacct accgtcacg gcgcggacgc	33960
gagtcggaat cggggggcgg ggtactgggc acgggtgccc accacgttca ccccgggta	34020
cgggcaggtg gaggcccat gaggctcgag atgttcgacc ccgccccgat cggcgctcgtg	34080
ttcaccagg ggccggagca ccggctcgcg tacaccaacg ccgtctaccg ggagaccttc	34140
ggcgaccgcc cgctggggcg gacgatccgc gaggccttcc ccgacctgcg gcagtccggc	34200
tacttcgaca tcttcgaccg ggtcctcacc acgggcgcgg ccgaggtggt caccgcggtg	34260
cccctcgacc tgatctacct cggctccacg ggcgagggca ggcgtactt cacgttcagc	34320
atctcccgcg ccacgatgag cgacggccgg ccgggagtg ctcggcgtgat cgtggaggtg	34380
accgcgcagg tgaccgcgcg ggaacggatc cgtgtgctgg ccgaggagcg ccgccgcgcg	34440
ctgcagcgct accgcagcct ggtgaacgcc ggaacgcaga tgggtgggt ggccgacgcc	34500
aaggccgga tcaccgagcc gagccccggc tgggaacgcg tgaccgggca gacctgggag	34560
gagttccgcg gcgagggtg gatgaacgcc gtccaccccg acgaccgcgc gcctcggtc	34620
gaggcgtggc ggccggcgac gaccgaacag gtgccgcgt ggatccacac ctaccggctg	34680
cggctggccg ccggcgggta ccggcacttc gtcgtcgacg ccgcgccgt gcgcgacggg	34740
aacacggtga tcgaatgggt gggcacctgc acggacatcg agcgggaatg gcaggagggc	34800
cgccgtacgg aactgctggc gcgggccgcc accgccacgt ccggcatcgc gcgctggac	34860
gagatgctcg ccgccctggc cgatgtgatc gtgcccgcgc tcgccgacaa ctgcaccatc	34920
cacctctgc cgcaggccct gcacctctg ccgggcaccc cgctgaccac cgaacgcgtc	34980

gcgcggtca cccgccggg gctcccggac ctgccccgc accacgagga gcacctgcgg	35040
cccggcagcc cgctggccc cgccgccgac cgccgcagcc cgctccactt cgtcttccc	35100
cccggcgagc cgccggccga cctcgctccg ctcgacggcg agccctggat ggccgaggac	35160
gtcaacagcg tcgtgctgct gcccgctgtc gtcgacggca ccaccgccgc cctggtcgcc	35220
gtctccacca gggcgccc cgcccccctc ggccaggcgg agatcggcct gctgcagaca	35280
ctcctggaac gcgccacac cccctcagc aacgccctgg agtaccagcg caccggcag	35340
gtggccctgg ccctgcagaa cagcctgctc accgaccgc cggacgcgc cggcctggac	35400
atgccgtcc gctaccggcc cagcaccgcc gccgccgagg tcggcgggga ctggtacgac	35460
gcgttcgtgc tgcgcgacgg cgccaccgtc ctcaccatcg gcgacgtctc cggccacgac	35520
ctgccggccg ccgtcaccat gagccagctg cgcaacatgc tgcgcgggct cacgctggac	35580
cgccaggaac cgaccggcac catcctgccc cggctggaca tcgccgtgca gacctctat	35640
acggagtgca ccgccacctg cgtgctggcc cgggtggaac gcccggaactc cggcggcgtc	35700
cggctgcact actccgtcgc cggtcaccgc ccgccgtgc tcgtcgaggc ggacggctcc	35760
gcgcgcttcc tgaccggggc gcggtccccg atgctcgggc tcgtccccgc gccggagtac	35820
tcgagcgcca tggaaccgct gccgccggc tccacctgc tgctgtacac cgacgggctg	35880
gtggagcgcc gcgacgagga tctcaccgtg ggctggagc ggctcggca ccacgcctcg	35940
gaggcggta gccgccgct gcaggacttc tgcgacacac tgctcaccgg ccagctcacc	36000
gtcgacaacg acgacgacgt ggcgatgctg gtcctgcgcc ggtaggagcg tgccgaggag	36060
cgccactctg gccgatttta cccttgcttt tccatcggga ttcgttctcc ggatttccc	36120
atccggcgcc gacggcgaga ccgttgggat caccaatacc ccggaattcc cgctccgcc	36180
accgttgggc agcgacggat cctgtgatat ttcgactacg cgcggtgatg aattggctcg	36240
gtgccggtcg cgccggctg tagcagttct ggagcgcgtc tggacatcgt cacgagcgt	36300
tgtgattctt ggtcctgtac acgcaagccg gcgcaacgtc cacgttccc atcagcggtt	36360

atcggcggtc caccggcgcg acggtgaccg cgggcgggta ctcatagggg gaactgcaat	36420
gaattactca aaagcagcga gaggaatgcc gacagccgga caaggtgccg ttogggcggc	36480
gcgcgtcgtc cgtgaaagtc cggcggaatc agaaacggtc acagttcaga tagcgtcgtt	36540
attaccgggt gagtcgtgc gtcgaaagg gatcgagcag aaccacgtcg cggcactcgc	36600
ggaggtagac gcgccgttc cggccatact ggtggaccgg aagacgatgc gggtcgtcga	36660
cgggatgcac cggctcctcg cggctctgct caacggacgg cagacgatcg aggccgaact	36720
gttcgacgga accgcggatg agggattcct gcgcgccgtc cgggagaacg tggcgcacgg	36780
actcccgtg tcgcaggcgg accgccgggc cgccgctcgc cgcatcatcg tgtcccacc	36840
gcattctgtc gacagggcga tcgccgggc gtccgggctc ggggcgaaga ccgtcgcggc	36900
cgtgcggcgc agttcaactg ccgtcgtgcc gcagttgaac acccgggtgg gccaggacgg	36960
cagggtccgg ccgctgaacg ggggcgaggg gcggcgcagg gccatggcgg tactggccga	37020
acaccccgac gcgtccctgc gcgaggtcgc ccgtctgtcc ggggtgtcgc ccgcgacggt	37080
cagcgacgta cgccggcggc tggccgccgg cgagtcgccc ctgccgtcga gacgggaacc	37140
ggccgaaccg cggacgggcg ccgactccca ccgcaaccag agcttcgttg atcccgctcc	37200
ggtgctggag aagctcgtgc gcgaccctc tctgcggcac aaggaggcg gccgccagct	37260
gtccagctg ctccgccaga acgcggtcgg cgtgcaggac ctgatggagc tgtccgacgc	37320
cgtgccgtcc cactgcaggt ccttggtgat ccatctcgcg cagcagtacc gggacgcctg	37380
gcagtccttc gcggagaagc tggacgagcc cgcctgcgcc tgtcccgggt gacgaacggg	37440
cggcacggac ccgttcaccg gacatgaccg gcgccgcgcc gcgttcacgg cgcgccgccg	37500
gcactcccac ggcacccgga ccaccgccgc gtatccggcg gaccggggcc cggcggggcc	37560
ggattcagcg ggcgggggcc cagggtgccac ccgattccag ccaccgggag agctccgccg	37620
ccgagtcctt gcgcacgacc agttcgacga ggccgcgggt ctggtcctga ccgtgctcga	37680
tgccgacgtc ctcgatgttg acgccaagt cgccgatcga cgtgaacagt tcggccaggg	37740
cgccgggctt gtcggagatg gtcaccgaga cggtcgcgag ctccgtccgg cgcgtaccgg	37800

gtttgcgcac gatcctggcg cacccccggt tcccctcccg caacagctcc tcgagctcct	37860
cctgcgcgcg gcggcggaacc agcgggtcgg cgtcggagac ggcgcgacgc gcgccgacgg	37920
cccgcccag gccggcggcg agggagtcga gaacgtccgc cacggccgtg gcgttggaac	37980
gcaggatgtc cccccagagc cgggcgtcac cggcccgat ccgggtgacg tcggcgacgc	38040
cctgccccgc cagccggacg ctgtcctccg ccgcgtgtc cagccgcgcg gcgagcaggg	38100
aggagagccg atggggcgcg tgcgagacga gggccaccgc gtggtcgtgc acaccggcgt	38160
ccatgaccac cggcattgcc tcgcacaacg acaccatctc caggcggtg ttcagcacgt	38220
cctgcccggg cagctccgac ggggtgagca ccagggcg cccctcgaag aggtccgcc	38280
gggcggcgag cggcccgaa cgtcgtgtgc cggccagcgg atggcttct atgtagctgg	38340
ccgggtcggc ccgcacgcg cgcacgtcgt cgtgcgggac cttcttgacg ctggcgacat	38400
cgaggtaggc tcgggccagc ccgtctcct gtgcgcgcgc gagcacgcgt ccgacctgtg	38460
ccgggggcac ggccagcacc gccaggtcga cctgacggtc cgtctctcc agggatccc	38520
cgcocatcgc ctccgccgtc ctggcggcgt tccggtcgac gtctccagg tgcacgccga	38580
ccccgcggcg ggtcagcgcg agagcgacgg acgtgccgat ggccccggtg ccgatgactg	38640
tggtgtcct caacgcgcgc cccaggtgc ggtgatccga aatcggctcg gacaagtgc	38700
gtgcccggca cgggaaaagg gaattccat ggcccggtgc gccccaatt taacgcttcg	38760
gcgcgatgt tcaactgcgg cgtgcagcg gtcgaacaca gtagcggtac accggacat	38820
tgaggcatcg tgctcagttg gcgacaccgg gtcggataaa cgcgggaatc cgaggagttg	38880
acgttgcagt cagcgtgag acacgacgac ctgcatccga tagaagaagt ggaaataagt	38940
togctctcca ccgacggctc ccgcgggatc gacggggaga gtcccagca cgtggaaatg	39000
ctggccgccg ccgacaccgc gcttcaccg atcatggtgc accccgcac cgggcgggtc	39060
atcgacggca tgcaccggct gcgcgccgcg atgctgacgg gccgtacgac gatcgcggtg	39120
aggttcttcg acggcaccca ggaggacgcc ttctcctcg ccgtgaagtc gaacatcgcg	39180

cacggactgc cgctgtccgc cgccgaccgc cggcggggccg ccggggcgcac catggccacc	39240
catccccggt ggtcggaccg gatgatcgcc tcggtggtcg gcacctccgc caggacggtc	39300
gccgagatcc gccgcgacgc cggcgccgcc ggggcggggg agcccacccg catcggccgg	39360
gacggcaggg tacggcccggt cgacgtgagc gagggccgca gactggccca cgacatgatc	39420
gtccgcgacc cgggcctgtc gctgcgccag gtcccccgcg ccgccgggat ctgccggag	39480
accgtcaggg acgtcagaca ccggatgctc cgcggtgagg acccggtgcc cgcgcccgcg	39540
ccgccgaccc tgggtggagcg cggcgccggac cgcggggcgg agccggccgg gaaggccgcc	39600
gcgccgtgcg ggacggagcc gccgcccgcg gtcgtgatga agcggctgag ggccgatccg	39660
gcgctgcgtc tcaacgagaa cggacgcgac ctgctgcggc ttctggatat ccacacggtc	39720
cggctggagg actggaaccg cattatcgaa agcgtgccgc cgcaccgtct ggagacggtg	39780
gcgagctgg cacgctcctg cgccgacaaa tggtcgaga tcgcgtcacg catcgaaagc	39840
aacgcatcac atctggccgg gtgaacgagg aaacacacga atccttcgag gagccgtcgg	39900
agaaagcggg acggcccgtc ggaacaccct tgttgagggg caatggagat acggtcgatc	39960
gatcacgtcg aattgttcgt cgaggacgcc caggacacgg ccggcaggct gtgcgactcc	40020
ttcggcttcg tccgcgtggg ccgcggcgcc gggaccaccg gactgcgcgg ctgcgagtcc	40080
gtcctgctgc gccagaacga catcgccctg ctgctgacca cggccaccga cgccgaccac	40140
cgtgccgccg agtacgtgaa gcagcacggg gacggggtcg cggatgatcg gtgggatccc	40200
cgggtaccga gctcgaattc gccctatagt gagtctgatt acaattcact ggccgtcgtt	40260
ttacaacgtc gtgactggga aaaccctggc gttaccaac ttaatgcct tgcagcacat	40320
cccccttcg ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttccaacag	40380
ttgcgcagct gaatggcgaa tggcgcccta tgcggtatit tctccttacg catctgtcg	40440
gtatttcaca ccgcatatgg tgcactctca gtacaatctg ctctgatgcc gcatagttaa	40500
gccagccccg acaccggcca acaccgctg acgcgaaccc cttgcggccg c	40551

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 127789-236560 (P005)	FOR FURTHER ACTION see Form PCT/ISA/220 as well as, where applicable, item 5 below.	
International application No. PCT/US2017/064328	International filing date (<i>day/month/year</i>) 1 December 2017 (01-12-2017)	(Earliest) Priority Date (<i>day/month/year</i>) 6 December 2016 (06-12-2016)
Applicant OREGON STATE UNIVERSITY		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 8 sheets.

☒ It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

a. With regard to the **language**, the international search was carried out on the basis of:

- ☒ the international application in the language in which it was filed
☐ a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))

b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6**bis**(a)).

c. ☒ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☐ **Certain claims were found unsearchable** (See Box No. II)

3. ☒ **Unity of invention is lacking** (see Box No III)

4. With regard to the **title**,

- ☒ the text is approved as submitted by the applicant
☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

- ☒ the text is approved as submitted by the applicant
☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority

6. With regard to the **drawings**,

- a. the figure of the **drawings** to be published with the abstract is Figure No. 1
☒ as suggested by the applicant
☐ as selected by this Authority, because the applicant failed to suggest a figure
☐ as selected by this Authority, because this figure better characterizes the invention
b. ☐ none of the figures is to be published with the abstract

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/064328

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
- a. ☒ forming part of the international application as filed:
- ☒ in the form of an Annex C/ST.25 text file.
- ☐ on paper or in the form of an image file.
- b. ☐ furnished together with the international application under PCT Rule 13~~ter~~.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
- ☐ in the form of an Annex C/ST.25 text file (Rule 13~~ter~~.1(a)).
- ☐ on paper or in the form of an image file (Rule 13~~ter~~.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/064328

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☒ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2017/064328

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/36 C07K11/02 C12P21/02
ADD. C12R1/465

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, COMPENDEX, EMBASE, FSTA, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2008/054945 A2 (OREGON STATE [US]; YIN XIHOU [US]; ZABRISKIE T MARK [US]) 8 May 2008 (2008-05-08) page 25, last paragraph; claims 10-12 -----	1-27
X	X. YIN: "The enduracidin biosynthetic gene cluster from Streptomyces fungicidicus", MICROBIOLOGY, vol. 152, no. 10, 1 October 2006 (2006-10-01), pages 2969-2983, XP055444228, GB ISSN: 1350-0872, DOI: 10.1099/mic.0.29043-0 the whole document ----- -/--	1-27



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

15 March 2018

Date of mailing of the international search report

29/03/2018

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Lejeune, Robert

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2017/064328

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BALTZ RICHARD H: "Genetic manipulation of secondary metabolite biosynthesis for improved production in Streptomyces and other actinomycetes", JOURNAL OF INDUSTRIAL MICROBIOLOGY AND BIOTECHNOLOGY, BASINGSTOKE, GB, vol. 43, no. 2, 12 September 2015 (2015-09-12), pages 343-370, XP035895568, ISSN: 1367-5435, DOI: 10.1007/S10295-015-1682-X [retrieved on 2015-09-12] page 346 - page 349 -----	1-27
A	CN 105 039 382 A (UNIV TIANJIN SCIENCE & TECH; TIANJIN XINXING VETERINARY PHARMACEUTICAL) 11 November 2015 (2015-11-11) the whole document -----	1-27
A	TIANJIE YUAN ET AL: "Improvement of antibiotic productivity by knock-out of in", MICROBIOLOGICAL RESEARCH, FISCHER, JENA, DE, vol. 166, no. 7, 26 October 2010 (2010-10-26), pages 539-547, XP028323848, ISSN: 0944-5013, DOI: 10.1016/J.MICRES.2010.10.006 [retrieved on 2010-11-18] the whole document -----	1-27
A	ANA YEPES ET AL: "Novel Two-Component Systems Implied in Antibiotic Production in Streptomyces coelicolor", PLOS ONE, vol. 6, no. 5, 20 May 2011 (2011-05-20), page e19980, XP055444246, DOI: 10.1371/journal.pone.0019980 the whole document -----	1-27
A	B. OSTASH ET AL: "Identification and characterization of the Streptomyces globisporus 1912 regulatory gene IndYR that affects sporulation and antibiotic production", MICROBIOLOGY, vol. 157, no. 4, 3 February 2011 (2011-02-03), pages 1240-1249, XP55444250, GB ISSN: 1350-0872, DOI: 10.1099/mic.0.045088-0 the whole document -----	1-27

	-/--	

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2017/064328

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SEON-HYE KIM ET AL: "Transcriptome Analysis of an Antibiotic Downregulator Mutant and Synergistic Actinorhodin Stimulation via Disruption of a Precursor Flux Regulator in Streptomyces coelicolor", APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 77, no. 5, 1 March 2011 (2011-03-01), pages 1872-1877, XP55444251, ISSN: 0099-2240, DOI: 10.1128/AEM.02346-10 the whole document -----	1-27
A	XIHOU YIN ET AL: "Enduracidin Analogues with Altered Halogenation Patterns Produced by Genetically Engineered Strains of Streptomyces fungicidicus", JOURNAL OF NATURAL PRODUCTS., vol. 73, no. 4, 23 April 2010 (2010-04-23) , pages 583-589, XP055444817, US ISSN: 0163-3864, DOI: 10.1021/np900710q the whole document -----	1-27
A	SEON-HYE KIM ET AL: "ABSTRACT", APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 77, no. 5, 1 March 2011 (2011-03-01), pages 1872-1877, XP055444251, ISSN: 0099-2240, DOI: 10.1128/AEM.02346-10 the whole document -----	1-27

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2017/064328

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2008054945 A2	08-05-2008	US 2010035256 A1 WO 2008054945 A2	11-02-2010 08-05-2008
CN 105039382 A	11-11-2015	CN 104561048 A CN 105039382 A	29-04-2015 11-11-2015

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 14, 22-26(completely); 1-13, 16, 17(partially)

A recombinant strain of *Streptomyces fungicidicus* comprising an augmented open reading frame-24 (orf24), wherein an enhanced production of enduracidin is obtained with the recombinant strain of *Streptomyces fungicidicus* in comparison to that obtained with a control *Streptomyces fungicidicus* strain.

2. claims: 15, 18-21, 27(completely); 1-13, 16, 17(partially)

A recombinant strain of *Streptomyces fungicidicus* comprising a diminished open reading frame-18 (orf18), wherein an enhanced production of enduracidin is obtained with the recombinant strain of *Streptomyces fungicidicus* in comparison to that obtained with a control *Streptomyces fungicidicus* strain.

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:

see form PCT/ISA/220

PCT

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY (PCT Rule 43bis.1)

Date of mailing
(day/month/year) see form PCT/ISA/210 (second sheet)

Applicant's or agent's file reference
see form PCT/ISA/220

FOR FURTHER ACTION
See paragraph 2 below

International application No. PCT/US2017/064328	International filing date (day/month/year) 01.12.2017	Priority date (day/month/year) 06.12.2016
--	--	--

International Patent Classification (IPC) or both national classification and IPC
INV. C07K14/36 C07K11/02 C12P21/02 ADD. C12R1/465

Applicant
OREGON STATE UNIVERSITY

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☒ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☐ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. FURTHER ACTION

If a demand for international preliminary examination is made, this opinion will usually be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1bis(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

Name and mailing address of the ISA:  European Patent Office P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk - Pays Bas Tel. +31 70 340 - 2040 Fax: +31 70 340 - 3016	Date of completion of this opinion see form PCT/ISA/210	Authorized Officer Lejeune, Robert Telephone No. +31 70 340-0	
---	---	--	---

Box No. I Basis of the opinion

1. With regard to the **language**, this opinion has been established on the basis of:
 - ☒ the international application in the language in which it was filed.
 - ☐ a translation of the international application into , which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1 (b)).
2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43bis.1(a))
3. ☒ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed:
 - ☒ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
4. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
5. Additional comments:

Box No. IV Lack of unity of invention

1. ☒ In response to the invitation (Form PCT/ISA/206) to pay additional fees, the applicant has, within the applicable time limit:
- ☒ paid additional fees
 - ☐ paid additional fees under protest and, where applicable, the protest fee
 - ☐ paid additional fees under protest but the applicable protest fee was not paid
 - ☐ not paid additional fees
2. ☐ This Authority found that the requirement of unity of invention is not complied with and chose not to invite the applicant to pay additional fees.
3. This Authority considers that the requirement of unity of invention in accordance with Rule 13.1, 13.2 and 13.3 is
- ☐ complied with
 - ☒ not complied with for the following reasons:
see separate sheet
4. Consequently, this report has been established in respect of the following parts of the international application:
- ☒ all parts.
 - ☐ the parts relating to claims Nos.

Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes: Claims	<u>1-21, 25-27</u>
	No: Claims	<u>22-24</u>
Inventive step (IS)	Yes: Claims	
	No: Claims	<u>1-27</u>
Industrial applicability (IA)	Yes: Claims	<u>1-27</u>
	No: Claims	

2. Citations and explanations

see separate sheet

Re Item IV

Lack of unity of invention

The present application deals with methods to increase the level of enduracidin production by *Streptomyces fungicidicus*. The methods include the augmentation of one the genes of the enduracidin biosynthetic gene cluster (orf24) and/or the reduction of another gene of the enduracidin biosynthetic gene cluster (orf18).

The gene cluster, and the sequences of each of the genes, is known from WO2008/054945. Overexpressing and/or reducing genes involved in secondary metabolite formation in *Streptomyces* is generally known (see pages 346-349 in Baltz et al. 2015 Journal of Industrial Microbiology and Biotechnology 43(2):343-370), and gene manipulation in *Streptomyces fungicidicus* in order to increase the level of enduracidin is known from CN105039382.

Consequently, the common feature of the claims, i.e. changing the level of one of genes from the enduracidin gene cluster in order to increase the production of enduracidin, is not considered to be a special technical feature in the sense of rule 13.2 PCT.

The inventions are:

Invention 1 (claims: 14, 22-26(completely); 1-13, 16, 17(partially))

A recombinant strain of *Streptomyces fungicidicus* comprising an augmented open reading frame-24 (orf24), wherein an enhanced production of enduracidin is obtained with the recombinant strain of *Streptomyces fungicidicus* in comparison to that obtained with a control *Streptomyces fungicidicus* strain.

Invention 2 (claims: 15, 18-21, 27(completely); 1-13, 16, 17(partially))

A recombinant strain of *Streptomyces fungicidicus* comprising a diminished open reading frame-18 (orf18), wherein an enhanced production of enduracidin is obtained with the recombinant strain of *Streptomyces fungicidicus* in comparison to that obtained with a control *Streptomyces fungicidicus* strain.

The present opinion concerns both inventions, since the applicant has paid an additional search fee.

Re Item V

Reasoned statement with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

Reference is made to the following documents:

- D1 WO 2008/054945 A2 (OREGON STATE [US]; YIN XIHOU [US]; ZABRISKIE T MARK [US]) 8 May 2008 (2008-05-08)
- D2 BALTZ RICHARD H: "Genetic manipulation of secondary metabolite biosynthesis for improved production in *Streptomyces* and other actinomycetes",
JOURNAL OF INDUSTRIAL MICROBIOLOGY AND BIOTECHNOLOGY,
BASINGSTOKE, GB,
vol. 43, no. 2, 12 September 2015 (2015-09-12), pages 343-370,
XP035895568,
ISSN: 1367-5435, DOI: 10.1007/S10295-015-1682-X
[retrieved on 2015-09-12]
- D3 CN 105 039 382 A (UNIV TIANJIN SCIENCE & TECH; TIANJIN XINXING VETERINARY PHARMACEUTICAL) 11 November 2015 (2015-11-11)
- D4 X. YIN: "The enduracidin biosynthetic gene cluster from *Streptomyces fungicidicus*",
MICROBIOLOGY,
vol. 152, no. 10, 1 October 2006 (2006-10-01), pages 2969-2983,
XP055444228,
GB
ISSN: 1350-0872, DOI: 10.1099/mic.0.29043-0

Invention 1

Novelty (Art 33(2) PCT)

D1 discloses (see claims 10-12) an expression vector comprising a nucleic acid encoding orf24 (SEQ ID NO:24 in D1). D1 also discloses the use of heterologous constitutive promoters therefor (see paragraph spanning pages 39 and 40). Consequently, the subject matter of claims 22-24 is not new.

The subject matter of claims 1-14, 16, 17, 25, 26 is new.

Inventive step (Art 33(3) PCT)

The closest prior art is D1, which discloses the gene cluster for the biosynthesis of enduracidin. D1 proposes to increase the the amount of enduracidin by using one of the genes selected from SEQ ID NO:22, 24, 41, 42 or 43 (see paragraph spanning pages 25 and 26). These genes are identifies as being involved in the regulation of antibiotic production (see page 29, lines 19-20).

The present application differs from D1 in that a specific gene is selected for augmentation, i.e. orf24 (=SEQ ID NO:24 in D1).

The present application shows that overexpression of orf24 increases the titer of enduracidin.

The skilled person, faced with the problem of increasing the production of enduracidin in *Streptomyces fungicidicus*, would have used the teaching of D1 and tested the effect of the disclosed regulatory genes on the enduracidin production. This corresponds to the general teaching in the art that secondary metabolite formation in *Streptomyces* can be improved by overexpressing and/or reducing regulatory genes (see D2). Moreover, genetic manipulation of *Streptomyces fungicidicus* is known to the skilled person: in D3 a point mutation in is introduced in *rpsL* in order to increase enramycin (=enduracidin) production in *Streptomyces fungicidicus*.

Consequently, since D1 has earmarked orf24 as a likely regulatory gene for the biosynthesis of enduracidin, and since overexpression of this type of gene in order to increase secondary metabolite formation is standard practice (D2), and in the

absence of problems requiring inventive action, one can only conclude that augmenting orf24 in *Streptomyces fungicidicus* in order to improve the production of enduracidin, does not involve an inventive step.

None of the claims are currently limited to the double presence of both an augmented orf24 and a reduced orf18. It is also noted that no experimental data concerning the effect of such a double modification is provided. In the absence of such data it is not possible to evaluate whether the provision of the double modification involves an inventive step. The effect of the reduction of orf18 on its own falls under the scope of invention 2.

Invention 2

Novelty (Art 33(2) PCT)

The prior art discloses orf18 as a member of the biosynthetic operon for enduracidin in *Streptomyces fungicidicus* (SEQ ID NO:18 in D1, orf18 in table 1 of D4). The prior art does not disclose a *Streptomyces fungicidicus* with an altered orf18 resulting in a lower level of orf18 and which produces enduracidin. The prior art does not disclose a disruption vector for orf18 (SEQ ID NO:37).

Consequently, the subject matter of claims 1-13, 15, 16, 17-21, 27 is new.

It should be noted that this novelty assessment is based on a narrow interpretation of the term "diminished". Replacement of this term with structurally distinguishable features will be required.

Inventive step (Art 33(3) PCT)

The closest prior art is D1, which discloses the gene cluster for the biosynthesis of enduracidin. D1 proposes to increase the amount of enduracidin by using one of the genes selected from SEQ ID NO:22, 24, 41, 42 or 43 (see paragraph spanning pages 25 and 26). These genes are identified as being involved in the regulation of antibiotic production (see page 29, lines 19-20).

The present application differs from D1 in that a specific gene is selected for reduction, i.e. orf18 (=SEQ ID NO:18 in D1).

The present application shows that disruption of orf18 increases the titer of enduracidin.

The skilled person, faced with the problem of increasing the production of enduracidin in *Streptomyces fungicidicus*, would have used the teaching of D1 and tested the effect of the regulatory genes on the enduracidin production. This corresponds to the general teaching in the art that secondary metabolite formation in *Streptomyces* can be improved by overexpressing and/or reducing regulatory genes (see D2). Moreover, genetic manipulation of *Streptomyces fungicidicus* is known to the skilled person: in D3 a point mutation in *rpsL* is introduced in order to increase enramycin (=enduracidin) production in *Streptomyces fungicidicus*.

Since orf18 is identified as a regulator (Table 1 of D4) and therefore earmarked as a likely regulatory gene for the biosynthesis of enduracidin, and since manipulation of this type of gene in order to increase secondary metabolite formation is standard practice (D2), and in the absence of problems requiring inventive action, one can only conclude that disrupting orf18 in *Streptomyces fungicidicus* in order to improve the production of enduracidin, does not involve an inventive step.

Further remarks: the deposits receipts (ATCC deposits) have not been provided. Consequently, it is not possible to confirm whether the deposits meet the requirements (e.g. the strict requirements of Rule 31 of the EPC concerning the identity of the depositor and the applicant).