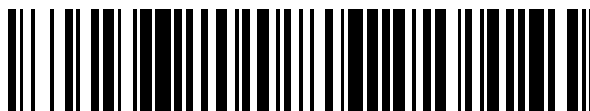


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 525 045**

51 Int. Cl.:

C12Q 1/68 (2006.01)
C12N 9/04 (2006.01)
C12N 9/20 (2006.01)
C12N 9/02 (2006.01)
C12N 9/10 (2006.01)
C12N 9/78 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.12.2008 E 08874034 (5)**

97 Fecha y número de publicación de la concesión europea: **27.08.2014 EP 2279265**

54 Título: **Agrupación de genes de saxitoxina de cianobacterias y detección de organismos cianotóxicos**

30 Prioridad:

24.04.2008 AU 2008902056

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

16.12.2014

73 Titular/es:

**NEWSOUTH INNOVATIONS PTY LIMITED
(100.0%)
Rupert Myers Building Barker Street University of
New South Wales
NSW 2052 , AU**

72 Inventor/es:

**NEILAN, BRETT, A.;
MIHALI, TROCO, KAAAN;
KELLMANN, RALF y
JEON, YOUNG, JAE**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 525 045 T3

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

DESCRIPCIÓN

Agrupación de genes de saxitoxina de cianobacterias y detección de organismos cianotóxicos

Campo Técnico

- 5 La presente invención se refiere a métodos para la detección de cianobacterias, dinoflagelados, y en particular, a métodos para la detección de organismos cianotóxicos. Se proporcionan kits para la detección de cianobacterias, dinoflagelados, y organismos cianotóxicos. La invención se refiere además a métodos para cribar para compuestos que modulan la actividad de polinucleótidos y/o polipéptidos de las rutas biosintéticas de la saxitoxina y cilindrospermopsina.

Antecedentes

- 10 Las cianobacterias, también conocidas como algas verdeazuladas, son bacterias fotosintéticas distribuidas ampliamente en entornos marinos y de agua dulce. Las cianobacterias que producen compuestos tóxicos son particularmente significativas para la calidad del agua y la salud humana y animal. En condiciones eutróficas, las cianobacterias tienden a formar grandes floraciones que promueven drásticamente concentraciones elevadas de toxinas. Las floraciones de cianobacterias pueden crecer mucho y expandirse en aguas costeras, arroyos, lagos y
15 agua potable y reservorios recreacionales. Las toxinas que producen pueden plantear un riesgo de salud grave para los seres humanos y animales y este problema es relevante internacionalmente ya que las cianobacterias más tóxicas tienen una distribución global.

- Es muy conocido un diverso rango de géneros de cianobacterias por la formación de floraciones de algas verdeazuladas tóxicas en superficies acuáticas. La saxitoxina (SXT) y sus análogos causan el síndrome de
20 intoxicación paralizante por moluscos (PSP), que aflige a la salud humana e impacta en las economías costeras de moluscos en todo el mundo. Las toxinas PSP son alcaloides únicos, que son producidas tanto por procariotas como eucariotas. Las toxinas PSP están entre las toxinas de algas más potentes y ubicuas y se consideran un riesgo toxicológico grave para la salud que puede afectar a los seres humanos, animales y ecosistemas en todo el mundo. Estas toxinas bloquean los canales de sodio y calcio regulados por voltaje y prolongan la apertura de los canales de
25 potasio evitando la transducción de señales neuronales. Se ha estimado que cada año se producen globalmente más de 2.000 casos humanos de PSP. Además, las floraciones costeras de microorganismos productores resulta en millones de dólares de daño económico debidos a la contaminación por la toxina PSP de marisco y el requerimiento continuo de programas costosos de monitorización de biotoxinas. Los sistemas de alerta temprana para anticipar las floraciones de algas productoras de la toxina paralizante de moluscos (PST), tales como PCR y cribado basado en
30 ELISA, todavía no están disponibles debido a la ausencia de datos sobre la base genética de la producción de PST.

- SXT es un alcaloide de perhidropurina tricíclica que puede estar sustituida en varias posiciones dando lugar a más de 30 análogos de SXT naturales. Aunque la biosíntesis de SXT parece compleja y única, los organismos de dos reinos, incluyendo determinadas especies de dinoflagelados marinos y cianobacterias de agua dulce, son capaces de producir estas toxinas, aparentemente por la misma ruta biosintética. A pesar de esfuerzos considerables, no se
35 ha identificado previamente ninguna de las enzimas o genes implicados en la biosíntesis y modificación de SXT.

- La aparición del género de cianobacterias *Cylindrospermopsis* se ha documentado en todos los continentes y por lo tanto plantea una amenaza significativa para la salud pública a escala global. La toxina principal producida por *Cylindrospermopsis* es cilindrospermopsina (CYR). Además de plantear una amenaza para la salud humana, la cilindrospermopsina también causa pérdidas económicas significativas para granjeros debido a la intoxicación del
40 ganado con agua potable contaminada por cilindrospermopsina. La cilindrospermopsina tiene efectos hepatotóxicos, citotóxicos y neurotóxicos generales y es un carcinógeno potencial. Su toxicidad se debe a la inhibición de la síntesis de glutatión y proteínas así como a la inhibición del citocromo P450. Hasta ahora se han identificado seis especies de cianobacterias que producen cilindrospermopsina; *Cylindrospermopsis raciborskii*, *Aphanizomenon ovalisporum*, *Aphanizomenon flos-aquae*, *Umezakia natans*, *Rhaphidiopsis curvata* y *Anabaena bergii*. Los incidentes de
45 intoxicación humana con cilindrospermopsina sólo se han reportado en Australia subtropical hasta la fecha, sin embargo *C. raciborskii* and *A. flos-aquae* se han detectado recientemente en áreas con climas más templados. La tendencia de *C. raciborskii* para formar floraciones densas y la capacidad de invasión de los organismos productores da lugar a preocupaciones globales respecto a la calidad del agua potable y necesita la monitorización de las reservas de agua potable para detectar la presencia de productores de cilindrospermopsina.

- 50 Existe una necesidad de métodos rápidos y exactos que detecten cianobacterias, y en particular aquellas cepas que son capaces de producir cianotoxinas tales como saxitoxina y cilindrospermopsina. Se necesitan métodos rápidos y exactos para detectar organismos cianotóxicos para evaluar el peligro potencial para la salud de floraciones de cianobacterias y para la implementación de estrategias de gestión del agua eficaces para minimizar los efectos de brotes de floraciones tóxicas.

Resumen

En un primer aspecto, se proporciona un polinucleótido aislado que comprende una secuencia según SEQ ID NO: 1 o una variante que comparte al menos 80% de identidad con SEQ ID NO: 1 o fragmento de ésta, en la que el

- fragmento comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68.
- En un segundo aspecto, se proporciona un ácido ribonucleico aislado o un ADN complementario aislado codificado por una secuencia según el primer aspecto.
- En un tercer aspecto, se proporciona un polipéptido aislado de la ruta biosintética de la saxitoxina codificado por el fragmento del primer aspecto y que comprende una secuencia de aminoácidos que comparte al menos 80% de identidad con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 69.
- En una realización, se proporciona una sonda o cebador que hibrida específicamente con uno o más de: un polinucleótido según el primer aspecto, un ácido ribonucleico o ADN complementario según el segundo aspecto, o un polipéptido según el tercer aspecto.
- Se describe un vector que comprende un polinucleótido según el primer aspecto, o un ácido ribonucleico o ADN complementario según el segundo aspecto. El vector puede ser un vector de expresión.
- Se describe una célula huésped que comprende el vector.
- En otra realización, se proporciona un anticuerpo aislado capaz de unirse específicamente a un polipéptido según el tercer aspecto.
- En un cuarto aspecto, se proporciona un método para detectar un organismo cianotóxico que comprende las etapas de obtener una muestra para usarse en el método y analizar la muestra para la presencia de una agrupación de genes SXT presente sólo en organismos productores de saxitoxina, en el que dicho análisis comprende detectar:
- (i) un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, y SEQ ID NO: 36;
 - (ii) un polinucleótido variante que tiene al menos 80% de identidad de secuencia con un polinucleótido de (i);
 - (iii) un ácido ribonucleico o ADN complementario codificado por una secuencia según (i);
 - (iv) un polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, y SEQ ID NO: 37; o,
 - (v) un polipéptido variante que tiene al menos 80% de identidad de secuencia con un polipéptido de (iv),
- y en el que dicha presencia es indicativa de organismos cianotóxicos en la muestra.
- En una realización del cuarto aspecto, el organismo cianotóxico es una cianobacteria o un dinoflagelado.
- En una realización del cuarto aspecto, el análisis de la muestra comprende la amplificación del ADN de la muestra por la reacción en cadena de la polimerasa y la detección de las secuencias amplificadas. La reacción en cadena de la polimerasa puede utilizar uno o más cebadores que comprenden una secuencia seleccionada del grupo que consiste en SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polinucleótido.
- En otra realización del cuarto aspecto, el método comprende analizar adicionalmente la muestra para la presencia de una o más de:
- (i) un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polinucleótido,

(ii) un ácido ribonucleico o ADN complementario codificado por una secuencia según (i),

(iii) un polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, y SEQ ID NO: 110 y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polipéptido.

El análisis adicional de la muestra puede comprender la amplificación del ADN de la muestra por la reacción en cadena de la polimerasa. La reacción en cadena de la polimerasa puede utilizar uno o más cebadores que comprenden una secuencia seleccionada del grupo que consiste en SEQ ID NO: 111, SEQ ID NO: 112, o variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polinucleótido.

Se describe un método para la detección de dinoflagelados, comprendiendo el método las etapas de obtener una muestra para usarse en el método y analizar la muestra para la presencia de uno o más de:

(i) un polinucleótido que comprende una secuencia según el primer aspecto,

(ii) un ácido ribonucleico o ADN complementario según el segundo aspecto,

(iii) un polipéptido que comprende una secuencia según el tercer aspecto,

en el que dicha presencia es indicativa de dinoflagelados en la muestra.

El análisis de la muestra puede comprender la amplificación del ADN de la muestra por la reacción en cadena de la polimerasa y la detección de las secuencias amplificadas.

En una realización del cuarto aspecto, la detección comprende una o ambas de electroforesis en gel y secuenciación de ácido nucleico. La muestra puede comprender uno o más organismos aislados o cultivados. La muestra puede ser una muestra medioambiental. La muestra medioambiental puede obtenerse de agua salada, agua dulce o una floración de algas verdeazuladas.

En un quinto aspecto, se proporciona un kit para la detección de organismos cianotóxicos, comprendiendo el kit al menos un agente para detectar la presencia de una agrupación de genes SXT presente sólo en organismos productores de saxitoxina, en el que bien:

(i) dicho agente puede detectar un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polinucleótido y dicho agente es un cebador o una sonda;

(ii) dicho agente puede detectar un ácido ribonucleico o ADN complementario codificado por una secuencia según (i) y dicho agente es un cebador o una sonda; o

(iii) dicho agente puede detectar un polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 37, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polipéptido, y dicho agente es un anticuerpo capaz de unirse específicamente al polipéptido.

El cebador o sonda puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134 y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polipéptido.

En otra realización del quinto aspecto, el kit comprende además al menos un agente adicional para detectar la presencia de uno o más de:

(i) un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polinucleótido,

(ii) un ácido ribonucleico o ADN complementario codificado por una secuencia según (i),

(iii) un polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, y SEQ ID NO: 110, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polipéptido.

El al menos un agente adicional puede ser un cebador, anticuerpo o sonda. El cebador o sonda puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 109, SEQ ID NO: 110, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polipéptido.

Se describe un kit para la detección de dinoflagelados, comprendiendo el kit al menos un agente para detectar la presencia de uno o más de:

(i) un polinucleótido que comprende una secuencia según el primer aspecto,

(ii) un ácido ribonucleico o ADN complementario según el segundo aspecto,

(iii) un polipéptido que comprende una secuencia según el tercer aspecto,

en el que dicha presencia es indicativa de dinoflagelados en la muestra.

Se describe además un método para cribar para un compuesto que modula la expresión o actividad de uno o más polipéptidos según el tercer aspecto, comprendiendo el método poner en contacto el polipéptido con un compuesto candidato en condiciones adecuadas para permitir la interacción del compuesto candidato y el polipéptido, y ensayar para actividad del polipéptido.

La modulación de la expresión o actividad de uno o más polipéptidos puede comprender inhibir la expresión o actividad de dicho polipéptido.

La modulación de la expresión o actividad de uno o más polipéptidos puede comprender aumentar la expresión o actividad de dicho polipéptido.

Definiciones

Tal y como se usa en esta solicitud, la forma singular "un", "una" y "el" incluyen las referencias plurales a no ser que el contexto dicte claramente otra cosa. Por ejemplo, el término "una célula madre" también incluye una pluralidad de células madre.

Tal y como se usa en la presente memoria, el término "que comprende" significa "que incluye." Las variaciones de la palabra "que comprende", tales como "comprenden" y "comprende," tienen correspondientemente significados variados. Así, por ejemplo, un polinucleótido "que comprende" una secuencia que codifica una proteína puede consistir exclusivamente en esa secuencia o puede incluir una o más secuencias adicionales.

Tal y como se usan en la presente memoria, los términos "anticuerpo" y "anticuerpos" incluyen IgG (incluyendo IgG1, IgG2, IgG3, e IgG4), IgA (incluyendo IgA1 e IgA2), IgD, IgE, o IgM, e IgY, anticuerpos completos, incluyendo anticuerpos completos de cadena única, y fragmentos de unión a antígeno de éstos. Los fragmentos de anticuerpo de unión a antígeno incluyen, pero no están limitados a, Fab, Fab' y F(ab')₂, Fd, Fvs de cadena única (scFv), anticuerpos de cadena única, Fvs unidos por disulfuro (sdFv) y fragmentos que comprenden bien un dominio VL o VH. Los anticuerpos pueden ser de cualquier origen animal. Los fragmentos de anticuerpo de unión a antígeno, incluyendo anticuerpos de cadena única, pueden comprender la o las regiones variables solas o en combinación con el total o parcial de lo siguiente: región bisagra, dominios CH1, CH2, y CH3. También se incluyen cualesquiera combinaciones de región o regiones variables y región bisagra, dominios CH1, CH2, y CH3. Los anticuerpos pueden ser monoclonales, policlonales, quiméricos, multiespecíficos, humanizados y anticuerpos humanos monoclonales y policlonales que se unen específicamente a la molécula biológica.

Tal y como se usan en la presente memoria, los términos "polipéptido" y "proteína" se usan indistintamente y se considera que tienen el mismo significado.

Tal y como se usan en la presente memoria, los términos "secuencia de nucleótidos" y "secuencia de polinucleótidos" se usan indistintamente y se considera que tienen el mismo significado.

Tal y como se usa en la presente memoria, el término "kit" se refiere a cualquier sistema de administración para administrar materiales. En el contexto de los ensayos de detección descritos en la presente memoria, dichos sistemas de administración incluyen sistemas que permiten el almacenamiento, transporte, o administración de reactivos de reacción (por ejemplo, etiquetas, muestras de referencia, material de soporte, etc. en los contenedores apropiados) y/o materiales de soporte (por ejemplo, tampones, instrucciones escritas para realizar el ensayo etc.) de una localización a otra. Por ejemplo, los kits incluyen uno o más cubículos, tales como cajas, que contienen los reactivos de reacción y/o materiales de soporte relevantes.

Cualquier discusión de documentos, actos, materiales, dispositivos, artículos o semejantes que se ha incluido en la presente especificación es solamente para el propósito de proporcionar un contexto para la presente invención. No debe considerarse como una admisión de que ninguna o todas estas materias forman parte de la base de la técnica anterior o formaban parte del conocimiento general común en el campo relevante a la presente invención antes de la fecha de prioridad de esta solicitud.

Descripción breve de los dibujos

Se describirá ahora una realización preferida de la presente invención, sólo como ejemplo, con referencia a los dibujos adjuntos en los que:

La Figura 1A es una tabla que muestra la distribución de los genes *sxt* en cianobacterias tóxicas y no tóxicas. PSP, saxitoxina; CYLN, cilindrospermopsina; +, fragmento génico amplificado; - no se detectó ningún gen.

La Figura 1B es una tabla que muestra secuencias de cebador usadas para amplificar varios genes SXT.

La Figura 2 es una tabla que muestra los genes *sxt* de la agrupación de genes de la saxitoxina de *C. raciborskii* T3, su longitud posible, su concordancia de similitud por BLAST con secuencias de proteínas similares de otros organismos, y su función predicha.

La Figura 3 es un diagrama que muestra la organización estructural de la agrupación de genes *sxt* de *C. raciborskii* T3. Las abreviaturas usadas son: IS4, secuencia de inserción 4; *at*, aminotransferasa; *dmt*, transportador de metabolito de fármaco; *ompR*, regulador transcripcional de la familia *ompR*; *penP*, unión a penicilina; *smf*, gen que se predice que está implicado en la captación de ADN. La escala indica la longitud de la agrupación de genes en pares de bases.

La Figura 4 es un diagrama de flujo que muestra la ruta para la biosíntesis de SXT y las funciones posibles de los genes *sxt*.

La Figura 5 muestra los espectros MS/MS de iones seleccionados de extractos celulares de *Cylindrospermopsis raciborskii* T3. Se indica la fragmentación predicha de iones y los valores *m/z* correspondientes. Figura 5A, arginina (*m/z* 175); Figura 5B, saxitoxina (*m/z* 300); Figura 5C, intermedio A' (*m/z* 187); Figura 5D, intermedio C' (*m/z* 211); Figura 5E, intermedio E' (*m/z* 225).

La Figura 6 es una tabla que muestra los genes *cyr* de la agrupación de genes de la cilindrospermopsina de *C. raciborskii* AWT205, su longitud posible, su concordancia de similitud por BLAST con secuencias de proteínas similares de otros organismos, y su función predicha.

La Figura 7 es una tabla que muestra la distribución del gen sulfotransferasa (*cyrJ*) en cianobacterias tóxicas y no tóxicas. La amplificación del gen de ARNt 16S se muestra como un control positivo. CYLN, cilindrospermopsina; SXT, saxitoxina; N.D., no detectado; +, fragmento génico amplificado; -, no se detectó ningún gen. NA, no disponible; AWQC, Australian Water Quality Center.

La Figura 8 es un diagrama de flujo que muestra la ruta biosintética de la biosíntesis de cilindrospermopsina.

La Figura 9 es un diagrama que muestra la organización estructural de la agrupación de genes de cilindrospermopsina de *C. raciborskii* AWT205. La escala indica la longitud de la agrupación de genes en pares de bases.

Descripción

Los inventores han identificado una agrupación de genes responsable de la biosíntesis de saxitoxina (la agrupación de genes SXT) y una agrupación de genes responsable de la biosíntesis de cilindrospermopsina (la agrupación de genes CYR). Se ha determinado la secuencia completa de cada agrupación de genes y se han identificado en ellas las actividades funcionales asignadas a cada uno de los genes. Tomando como base esta información, los inventores han elucidado las rutas biosintéticas completas de la saxitoxina y cilindrospermopsina.

De acuerdo con esto, la invención proporciona secuencias de polinucleótido y polipéptido derivadas de cada una de las agrupaciones de genes SXT y CYR y en particular, las secuencias relacionadas con los genes específicos en cada ruta. Se proporcionan métodos y kits para la detección de cepas de cianobacterias en una muestra tomando como base la presencia (o ausencia) en la muestra de una o más de las secuencias de la invención. Los inventores han determinado que determinados marcos de lectura abierta presentes en la agrupación de genes SXT de microorganismos productores de saxitoxina están ausentes en la agrupación de genes SXT de microorganismos que no producen saxitoxina. De manera similar, se ha descubierto que un marco de lectura abierto presente en la agrupación de genes CYR de microorganismos productores de cilindrospermopsina está ausente en microorganismos que no son productores de cilindrospermopsina. De acuerdo con esto, la invención proporciona métodos y kits para la detección de microorganismos productores de toxina.

También se proporcionan por la invención métodos de cribado para la identificación de compuestos capaces de modular la expresión o actividad de proteínas en las rutas biosintéticas de la saxitoxina y/o cilindrospermopsina.

Polinucleótidos y polipéptidos

5 Los inventores han determinado la secuencia de polinucleótido completa de la agrupación de genes de la saxitoxina (SXT) y la agrupación de genes de la cilindrospermopsina (CYR).

Según aspectos y realizaciones de la invención, la agrupación de genes SXT puede tener, pero no está limitado a, la secuencia de polinucleótido como se muestra en SEQ ID NO: 1 (número de registro GenBank DQ787200), o presentar una identidad de secuencia suficiente con ésta para hibridar con la secuencia de SEQ ID NO: 1.

La agrupación de genes SXT comprende 31 genes y 30 regiones intergénicas.

10 El Gen 1 de la agrupación de genes SXT es una secuencia de nucleótidos de 759 pares de bases (pb) mostrada en SEQ ID NO: 4. La secuencia de nucleótidos del Gen 1 de SXT abarca desde el nucleótido en la posición 1625 hasta el nucleótido en la posición 2383 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 1 (SXTD) se muestra en SEQ ID NO: 5.

15 El Gen 2 de la agrupación de genes SXT es una secuencia de nucleótidos de 396 pb mostrada en SEQ ID NO: 6. La secuencia de nucleótidos del Gen 2 de SXT abarca desde el nucleótido en la posición 2621 hasta el nucleótido en la posición 3016 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 2 (ORF3) se muestra en SEQ ID NO: 7.

20 El Gen 3 de la agrupación de genes SXT es una secuencia de nucleótidos de 360 pb mostrada en SEQ ID NO: 8. La secuencia de nucleótidos del Gen 3 de SXT abarca desde el nucleótido en la posición 2955 hasta el nucleótido en la posición 3314 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 3 (ORF4) se muestra en SEQ ID NO: 9.

25 El Gen 4 de la agrupación de genes SXT es una secuencia de nucleótidos de 354 pb mostrada en SEQ ID NO: 10. La secuencia de nucleótidos del Gen 4 de SXT abarca desde el nucleótido en la posición 3647 hasta el nucleótido en la posición 4000 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 4 (SXTC) se muestra en SEQ ID NO: 11.

El Gen 5 de la agrupación de genes SXT es una secuencia de nucleótidos de 957 pb mostrada en SEQ ID NO: 12. La secuencia de nucleótidos del Gen 5 de SXT abarca desde el nucleótido en la posición 4030 hasta el nucleótido en la posición 4986 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 5 (SXTB) se muestra en SEQ ID NO: 13.

30 El Gen 6 de la agrupación de genes SXT es una secuencia de nucleótidos de 3738 pb mostrada en SEQ ID NO: 14. La secuencia de nucleótidos del Gen 6 de SXT abarca desde el nucleótido en la posición 5047 hasta el nucleótido en la posición 8784 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 6 (SXTA) se muestra en SEQ ID NO: 15.

35 El Gen 7 de la agrupación de genes SXT es una secuencia de nucleótidos de 387 pb mostrada en SEQ ID NO: 16. La secuencia de nucleótidos del Gen 7 de SXT abarca desde el nucleótido en la posición 9140 hasta el nucleótido en la posición 9526 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 7 (SXTE) se muestra en SEQ ID NO: 17.

40 El Gen 8 de la agrupación de genes SXT es una secuencia de nucleótidos de 1416 pb mostrada en SEQ ID NO: 18. La secuencia de nucleótidos del Gen 8 de SXT abarca desde el nucleótido en la posición 9686 hasta el nucleótido en la posición 11101 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 8 (SXTF) se muestra en SEQ ID NO: 19.

45 El Gen 9 de la agrupación de genes SXT es una secuencia de nucleótidos de 1134 pb mostrada en SEQ ID NO: 20. La secuencia de nucleótidos del Gen 9 de SXT abarca desde el nucleótido en la posición 11112 hasta el nucleótido en la posición 12245 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 9 de SXT (SXTG) se muestra en SEQ ID NO: 21.

El Gen 10 de la agrupación de genes SXT es una secuencia de nucleótidos de 1005 pb mostrada en SEQ ID NO: 22. La secuencia de nucleótidos del Gen 10 de SXT abarca desde el nucleótido en la posición 12314 hasta el nucleótido en la posición 13318 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 10 (SXTH) se muestra en SEQ ID NO: 23.

50 El Gen 11 de la agrupación de genes SXT es una secuencia de nucleótidos de 1839 pb mostrada en SEQ ID NO: 24. La secuencia de nucleótidos del Gen 11 de SXT abarca desde el nucleótido en la posición 13476 hasta el nucleótido en la posición 15314 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 11 (SXTI) se muestra en SEQ ID NO: 25.

- El Gen 12 de la agrupación de genes SXT es una secuencia de nucleótidos de 444 pb mostrada en SEQ ID NO: 26. La secuencia de nucleótidos del Gen 12 de SXT abarca desde el nucleótido en la posición 15318 hasta el nucleótido en la posición 15761 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 12 (SXTJ) se muestra en SEQ ID NO: 27.
- 5 El Gen 13 de la agrupación de genes SXT es una secuencia de nucleótidos de 165 pb mostrada en SEQ ID NO: 28. La secuencia de nucleótidos del Gen 13 de SXT abarca desde el nucleótido en la posición 15761 hasta el nucleótido en la posición 15925 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 13 (SXTK) se muestra en SEQ ID NO: 29.
- 10 El Gen 14 de la agrupación de genes SXT es una secuencia de nucleótidos de 1299 pb mostrada en SEQ ID NO: 30. La secuencia de nucleótidos del Gen 14 de SXT abarca desde el nucleótido en la posición 15937 hasta el nucleótido en la posición 17235 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 14 (SXTL) se muestra en SEQ ID NO: 31.
- 15 El Gen 15 de la agrupación de genes SXT es una secuencia de nucleótidos de 1449 pb mostrada en SEQ ID NO: 32. La secuencia de nucleótidos del Gen 15 de SXT abarca desde el nucleótido en la posición 17323 hasta el nucleótido en la posición 18771 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 16 (SXTM) se muestra en SEQ ID NO: 33.
- 20 El Gen 16 de la agrupación de genes SXT es una secuencia de nucleótidos de 831 pb mostrada en SEQ ID NO: 34. La secuencia de nucleótidos del Gen 16 de SXT abarca desde el nucleótido en la posición 19119 hasta el nucleótido en la posición 19949 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 16 (SXTN) se muestra en SEQ ID NO: 35.
- El Gen 17 de la agrupación de genes SXT es una secuencia de nucleótidos de 774 pb mostrada en SEQ ID NO: 36. La secuencia de nucleótidos del Gen 17 de SXT abarca desde el nucleótido en la posición 20238 hasta el nucleótido en la posición 21011 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 17 (SXTX) se muestra en SEQ ID NO: 37.
- 25 El Gen 18 de la agrupación de genes SXT es una secuencia de nucleótidos de 327 pb mostrada en SEQ ID NO: 38. La secuencia de nucleótidos del Gen 18 de SXT abarca desde el nucleótido en la posición 21175 hasta el nucleótido en la posición 21501 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 18 (SXTW) se muestra en SEQ ID NO: 39.
- 30 El Gen 19 de la agrupación de genes SXT es una secuencia de nucleótidos de 1653 pb mostrada en SEQ ID NO: 40. La secuencia de nucleótidos del Gen 219 de SXT abarca desde el nucleótido en la posición 21542 hasta el nucleótido en la posición 23194 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 19 (SXTV) se muestra en SEQ ID NO: 41.
- 35 El Gen 20 de la agrupación de genes SXT es una secuencia de nucleótidos de 750 pb mostrada en SEQ ID NO: 42. La secuencia de nucleótidos del Gen 20 de SXT abarca desde el nucleótido en la posición 23199 hasta el nucleótido en la posición 23948 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 20 (SXTU) se muestra en SEQ ID NO: 43.
- 40 El Gen 21 de la agrupación de genes SXT es una secuencia de nucleótidos de 1005 pb mostrada en SEQ ID NO: 44. La secuencia de nucleótidos del Gen 21 de SXT abarca desde el nucleótido en la posición 24091 hasta el nucleótido en la posición 25095 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 21 (SXTT) se muestra en SEQ ID NO: 45.
- El Gen 22 de la agrupación de genes SXT es una secuencia de nucleótidos de 726 pb mostrada en SEQ ID NO: 46. La secuencia de nucleótidos del Gen 22 de SXT abarca desde el nucleótido en la posición 25173 hasta el nucleótido en la posición 25898 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 22 (SXTS) se muestra en SEQ ID NO: 47.
- 45 El Gen 23 de la agrupación de genes SXT es una secuencia de nucleótidos de 576 pb mostrada en SEQ ID NO: 48. La secuencia de nucleótidos del Gen 23 de SXT abarca desde el nucleótido en la posición 25974 hasta el nucleótido en la posición 26549 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 23 (ORF24) se muestra en SEQ ID NO: 49.
- 50 El Gen 24 de la agrupación de genes SXT es una secuencia de nucleótidos de 777 pb mostrada en SEQ ID NO: 50. La secuencia de nucleótidos del Gen 24 de SXT abarca desde el nucleótido en la posición 26605 hasta el nucleótido en la posición 27381 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 24 (SXTR) se muestra en SEQ ID NO: 51.
- El Gen 25 de la agrupación de genes SXT es una secuencia de nucleótidos de 777 pb mostrada en SEQ ID NO: 52. La secuencia de nucleótidos del Gen 25 de SXT abarca desde el nucleótido en la posición 27392 hasta el nucleótido

en la posición 28168 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 25 (SXTQ) se muestra en SEQ ID NO: 53.

El Gen 26 de la agrupación de genes SXT es una secuencia de nucleótidos de 1227 pb mostrada en SEQ ID NO: 54. La secuencia de nucleótidos del Gen 26 de SXT abarca desde el nucleótido en la posición 28281 hasta el nucleótido en la posición 29507 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 26 (SXTTP) se muestra en SEQ ID NO: 55.

El Gen 27 de la agrupación de genes SXT es una secuencia de nucleótidos de 603 pb mostrada en SEQ ID NO: 56. La secuencia de nucleótidos del Gen 27 de SXT abarca desde el nucleótido en la posición 29667 hasta el nucleótido en la posición 30269 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 27 (SXTQ) se muestra en SEQ ID NO: 57.

El Gen 28 de la agrupación de genes SXT es una secuencia de nucleótidos de 1350 pb mostrada en SEQ ID NO: 58. La secuencia de nucleótidos del Gen 28 de SXT abarca desde el nucleótido en la posición 30612 hasta el nucleótido en la posición 31961 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 28 (ORF29) se muestra en SEQ ID NO: 59.

El Gen 29 de la agrupación de genes SXT es una secuencia de nucleótidos de 666 pb mostrada en SEQ ID NO: 60. La secuencia de nucleótidos del Gen 29 de SXT abarca desde el nucleótido en la posición 32612 hasta el nucleótido en la posición 33277 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 29 (SXTY) se muestra en SEQ ID NO: 61.

El Gen 30 de la agrupación de genes SXT es una secuencia de nucleótidos de 1353 pb mostrada en SEQ ID NO: 62. La secuencia de nucleótidos del Gen 30 de SXT abarca desde el nucleótido en la posición 33325 hasta el nucleótido en la posición 34677 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 30 (SXTZ) se muestra en SEQ ID NO: 63.

El Gen 31 de la agrupación de genes SXT es una secuencia de nucleótidos de 819 pb mostrada en SEQ ID NO: 64. La secuencia de nucleótidos del Gen 31 de SXT abarca desde el nucleótido en la posición 35029 hasta el nucleótido en la posición 35847 de SEQ ID NO: 1. La secuencia de polipéptido codificada por el Gen 31 (OMPR) se muestra en SEQ ID NO: 65.

La región del extremo 5' de la agrupación de genes SXT comprende un gen de 1320 pb (orf1), cuya secuencia se muestra en SEQ ID NO: 2. La secuencia de nucleótidos de orf1 abarca desde el nucleótido en la posición 1 hasta el nucleótido en la posición 1320 de SEQ ID NO: 1. La secuencia de polipéptido codificada por orf1 se muestra en SEQ ID NO: 3.

La región del extremo 3' de la agrupación de genes SXT comprende un gen de 774 pb (hisA), cuya secuencia se muestra en SEQ ID NO: 66. La secuencia de nucleótidos de hisA abarca desde el nucleótido en la posición 35972 hasta el nucleótido en la posición 36745 de SEQ ID NO: 1. La secuencia de polipéptido codificada por hisA se muestra en SEQ ID NO: 67.

La región del extremo 3' de la agrupación de genes SXT también comprende un gen de 396 pb (orfA), cuya secuencia se muestra en SEQ ID NO: 68. La secuencia de nucleótidos de orfA abarca desde el nucleótido en la posición 37060 hasta el nucleótido en la posición 37455 de SEQ ID NO: 1. La secuencia de polipéptido codificada por orfA se muestra en SEQ ID NO: 69.

Según otros aspectos y realizaciones de la invención, la agrupación de genes CYR puede tener, pero no está limitado a, la secuencia de nucleótidos como se muestra en SEQ ID NO: 80 (número de registro GenBank EU140798), o presentar una identidad de secuencia suficiente con ésta para hibridar con la secuencia de SEQ ID NO: 80.

La agrupación de genes CYR comprende 15 genes y 14 regiones intergénicas.

El Gen 1 de la agrupación de genes CYR es una secuencia de nucleótidos de 5631 pb mostrada en SEQ ID NO: 81. La secuencia de nucleótidos del Gen 1 de CYR abarca desde el nucleótido en la posición 444 hasta el nucleótido en la posición 6074 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 1 (CYRD) se muestra en SEQ ID NO: 82.

El Gen 2 de la agrupación de genes CYR es una secuencia de nucleótidos de 4074 pb mostrada en SEQ ID NO: 83. La secuencia de nucleótidos del Gen 2 de CYR abarca desde el nucleótido en la posición 6130 hasta el nucleótido en la posición 10203 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 2 (CYRF) se muestra en SEQ ID NO: 84.

El Gen 3 de la agrupación de genes CYR es una secuencia de nucleótidos de 1437 pb mostrada en SEQ ID NO: 85. La secuencia de nucleótidos del Gen 3 de CYR abarca desde el nucleótido en la posición 10251 hasta el nucleótido

en la posición 11687 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 3 (CYRG) se muestra en SEQ ID NO: 86.

El Gen 4 de la agrupación de genes CYR es una secuencia de nucleótidos de 831 pb mostrada en SEQ ID NO: 87. La secuencia de nucleótidos del Gen 4 de CYR abarca desde el nucleótido en la posición 11741 hasta el nucleótido en la posición 12571 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 4 (CYRI) se muestra en SEQ ID NO: 88.

El Gen 5 de la agrupación de genes CYR es una secuencia de nucleótidos de 1398 pb mostrada en SEQ ID NO: 89. La secuencia de nucleótidos del Gen 5 de CYR abarca desde el nucleótido en la posición 12568 hasta el nucleótido en la posición 13965 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 5 (CYRK) se muestra en SEQ ID NO: 90.

El Gen 6 de la agrupación de genes CYR es una secuencia de nucleótidos de 750 pb mostrada en SEQ ID NO: 91. La secuencia de nucleótidos del Gen 6 de CYR abarca desde el nucleótido en la posición 14037 hasta el nucleótido en la posición 14786 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 6 (CYRL) se muestra en SEQ ID NO: 92.

El Gen 7 de la agrupación de genes CYR es una secuencia de nucleótidos de 1431 pb mostrada en SEQ ID NO: 93. La secuencia de nucleótidos del Gen 7 de CYR abarca desde el nucleótido en la posición 14886 hasta el nucleótido en la posición 16316 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 7 (CYRH) se muestra en SEQ ID NO: 94.

El Gen 8 de la agrupación de genes CYR es una secuencia de nucleótidos de 780 pb mostrada en SEQ ID NO: 95. La secuencia de nucleótidos del Gen 8 de CYR abarca desde el nucleótido en la posición 16893 hasta el nucleótido en la posición 17672 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 8 (CYRJ) se muestra en SEQ ID NO: 96.

El Gen 9 de la agrupación de genes CYR es una secuencia de nucleótidos de 1176 pb mostrada en SEQ ID NO: 97. La secuencia de nucleótidos del Gen 9 de CYR abarca desde el nucleótido en la posición 18113 hasta el nucleótido en la posición 19288 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 9 (CYRA) se muestra en SEQ ID NO: 98.

El Gen 10 de la agrupación de genes CYR es una secuencia de nucleótidos de 8754 pb mostrada en SEQ ID NO: 99. La secuencia de nucleótidos del Gen 10 de CYR abarca desde el nucleótido en la posición 19303 hasta el nucleótido en la posición 28056 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 10 (CYRB) se muestra en SEQ ID NO: 100.

El Gen 11 de la agrupación de genes CYR es una secuencia de nucleótidos de 5667 pb mostrada en SEQ ID NO: 101. La secuencia de nucleótidos del Gen 11 de CYR abarca desde el nucleótido en la posición 28061 hasta el nucleótido en la posición 33727 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 11 (CYRE) se muestra en SEQ ID NO: 102.

El Gen 12 de la agrupación de genes CYR es una secuencia de nucleótidos de 5004 pb mostrada en SEQ ID NO: 103. La secuencia de nucleótidos del Gen 12 de CYR abarca desde el nucleótido en la posición 34299 hasta el nucleótido en la posición 39302 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 12 (CYRC) se muestra en SEQ ID NO: 104.

El Gen 13 de la agrupación de genes CYR es una secuencia de nucleótidos de 318 pb mostrada en SEQ ID NO: 105. La secuencia de nucleótidos del Gen 13 de CYR abarca desde el nucleótido en la posición 39366 hasta el nucleótido en la posición 39683 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 13 (CYRM) se muestra en SEQ ID NO: 106.

El Gen 14 de la agrupación de genes CYR es una secuencia de nucleótidos de 600 pb mostrada en SEQ ID NO: 107. La secuencia de nucleótidos del Gen 14 de CYR abarca desde el nucleótido en la posición 39793 hasta el nucleótido en la posición 40392 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 14 (CYRN) se muestra en SEQ ID NO: 108.

El Gen 15 de la agrupación de genes CYR es una secuencia de nucleótidos de 1548 pb mostrada en SEQ ID NO: 109. La secuencia de nucleótidos del Gen 15 de CYR abarca desde el nucleótido en la posición 40501 hasta el nucleótido en la posición 42048 de SEQ ID NO: 80. La secuencia de polipéptido codificada por el Gen 15 (CYRO) se muestra en SEQ ID NO: 110.

En general, los ácidos nucleicos y polipéptidos de la invención están en una forma aislada o purificada.

Los polinucleótidos SXT y CYR descritos en la presente memoria pueden ser ácidos desoxirribonucleicos (ADN), ácidos ribonucleicos (ARN) o ácidos desoxirribonucleicos complementarios (ADNc).

El ARN puede derivar de la transcripción catalizada por una ARN polimerasa de una secuencia de ADN. El ARN puede ser un transcrito primario derivado de la transcripción de una secuencia de ADN correspondiente. El ARN también puede experimentar procesamiento posterior a la transcripción. Por ejemplo, un transcrito de ARN primario puede experimentar procesamiento posterior a la transcripción para formar un ARN maduro. ARN mensajero (ARNm) se refiere a ARN derivado de un marco de lectura abierto correspondiente que puede traducirse en proteína por la célula. ADNc se refiere a un ADN bicatenario que es complementario a y deriva de ARNm. ARN con sentido se refiere a transcrito de ARN que incluye el ARNm y así puede traducirse en proteína por la célula. ARN antisentido se refiere a un transcrito de ARN que es complementario a todo o parte de un transcrito primario diana o ARNm y puede usarse para bloquear la expresión de un gen diana.

El experto en la técnica reconocerá qué secuencias de ARN o ADNc codificadas por las secuencias de ADN de SXT y CYR descritas en la presente memoria pueden derivarse usando el código genético. Una secuencia de ARN puede derivar de una secuencia de ADN dada generando una secuencia que es complementaria a la secuencia de ADN particular. La secuencia complementaria puede generarse convirtiendo cada base citosina ('C') en la secuencia de ADN en una base guanina ('G'), cada base guanina ('G') en la secuencia de ADN en una base citosina ('C'), cada base timidina ('T') en la secuencia de ADN en una base adenina ('A'), y cada base adenina ('A') en la secuencia de ADN en una base uracilo ('U').

Una secuencia de ADN complementario (ADNc) puede derivar de una secuencia de ADN obteniendo una secuencia de ARN de la secuencia de ADN como anteriormente, y convirtiendo la secuencia de ARN en una secuencia de ADNc. Una secuencia de ARN puede convertirse en una secuencia de ADNc convirtiendo cada base citosina ('C') en la secuencia de ARN en una base guanina ('G'), cada base guanina ('G') en la secuencia de ARN en una base citosina ('C'), cada base uracilo ('U') en la secuencia de ARN en una base adenina ('A'), y cada base adenina ('A') en la secuencia de ARN en una base timidina ('T').

El término "variante" tal y como se usa en la presente memoria se refiere a una secuencia sustancialmente similar. En general, dos secuencias son "sustancialmente similares" si las dos secuencias tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son los mismos (porcentaje de "identidad de secuencia"), sobre una región especificada o, cuando no se especifica, sobre la secuencia completa. De acuerdo con esto, una "variante" de una secuencia de polinucleótido y polipéptido descrita en la presente memoria puede compartir al menos 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 83%, 85%, 88%, 90%, 93%, 95%, 96%, 97%, 98% ó 99% de identidad de secuencia con la secuencia de referencia.

En general, las variantes de secuencias de polipéptido poseen una actividad biológica cualitativa en común. Las variantes de secuencias de polinucleótidos codifican generalmente polipéptidos que poseen generalmente una actividad biológica cualitativa en común. También se incluyen en el significado del término "variante" los homólogos de polinucleótidos y polipéptidos de la invención. Un homólogo de polinucleótido es típicamente de una especie bacteriana diferente pero comparte sustancialmente la misma función o actividad biológica que el polinucleótido correspondiente descrito en la presente memoria. Un homólogo de polipéptido es típicamente de una especie bacteriana diferente pero comparte sustancialmente la misma función o actividad biológica que el polipéptido correspondiente descrito en la presente memoria. Por ejemplo, los homólogos de los polinucleótidos y polipéptidos descritos en la presente memoria incluyen, pero no se limitan a, aquellos de diferentes especies de cianobacterias.

Además, el término "variante" también incluye análogos de los polipéptidos de la invención. Un "análogo" de polipéptido es un polipéptido que es un derivado de un polipéptido de la invención, derivado que comprende la adición, delección, sustitución de uno o más aminoácidos, de manera que el polipéptido retiene sustancialmente la misma función. El término "sustitución de aminoácidos conservativa" se refiere a una sustitución o reemplazo de un aminoácido por otro aminoácido con propiedades similares en una cadena de polipéptido (secuencia primaria de una proteína). Por ejemplo, la sustitución del aminoácido cargado ácido glutámico (Glu) por el aminoácido cargado similarmente ácido aspártico (Asp) sería una sustitución de aminoácidos conservativa.

En general, el porcentaje de identidad de secuencia entre dos secuencias puede determinarse comparando dos secuencias alineadas de forma óptima sobre una ventana de comparación. La parte de la secuencia en la ventana de comparación puede comprender, por ejemplo, delecciones o adiciones (es decir, huecos) en comparación con la secuencia de referencia (por ejemplo, una secuencia de polinucleótido o polipéptido descrita en la presente memoria), que no comprende delecciones o adiciones, con el fin de alinear las dos secuencias de forma óptima. Un porcentaje de identidad de secuencia puede calcularse entonces determinando el número de posiciones en las que aparecen bases de ácidos nucleicos o residuos de aminoácidos idénticos en ambas secuencias para rendir el número de posiciones concordantes, dividiendo el número de posiciones concordantes por el número total de posiciones en la ventana de comparación, y multiplicando el resultado por 100 para rendir el porcentaje de identidad de secuencia.

En el contexto de dos o más secuencias de ácido nucleico o polipéptido, el porcentaje de identidad de secuencia se refiere al porcentaje especificado de residuos de aminoácidos o nucleótidos que son iguales sobre la región especificada, (o, cuando no se especifica, sobre la secuencia completa), cuando se comparan y alinean para correspondencia máxima sobre una ventana de comparación, o región designada según se mide usando uno de los algoritmos de comparación de secuencias siguientes o por alineamiento manual e inspección visual.

Para comparaciones de secuencias, típicamente una secuencia actúa como una secuencia de referencia, con la que se comparan las secuencias de ensayo. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de ensayo y de referencia se introdujeron en un ordenador, se designan coordenadas de subsecuencia, si es necesario, y se designan los parámetros del programa del algoritmo de secuencia. Pueden usarse los parámetros del programa por defecto, o pueden designarse parámetros alternativos. El algoritmo de comparación de secuencias calcula el porcentaje de identidades de secuencia para las secuencias de ensayo respecto a la secuencia de referencia, tomando como base los parámetros del programa. Los métodos para el alineamiento de secuencias para comparación son muy conocidos en la técnica. El alineamiento óptimo de secuencias para comparación puede determinarse convencionalmente usando programas informáticos conocidos, incluyendo, pero no limitado a: CLUSTAL en el programa PC/Gene (disponible en Intelligenetics, Mountain View, California); el programa ALIGN (Versión 2.0) y GAP, BESTFIT, BLAST, FASTA, y TFASTA en el Paquete de Software GCG Wisconsin Genetics, Versión 10 (disponible en Accelrys Inc., 9685 Scranton Road, San Diego, California, EEUU).

El programa BESTFIT (Wisconsin Sequence Analysis Package, para Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711) usa el algoritmo de homología local de Smith y Waterman para encontrar el mejor segmento de homología entre dos secuencias (Smith y Waterman, *Advances in Applied Mathematics* 2:482-489 (1981)). Cuando se usa BESTFIT o cualquier otro programa de alineamiento de secuencias para determinar el grado de homología entre secuencias, los parámetros pueden ajustarse de manera que el porcentaje de identidad se calcula sobre la longitud completa de la secuencia de nucleótidos de referencia y se permiten los huecos en homología de hasta 5% del número total de nucleótidos en la secuencia de referencia.

GAP usa el algoritmo descrito en Needleman y Wunsch (1970) *J. Mol. Biol.* 48:443-453, para encontrar el alineamiento de dos secuencias completas que maximiza el número de concordancias y minimiza el número de huecos. GAP considera todos los alineamientos y posiciones de hueco posibles y crea el alineamiento con el mayor número de bases concordantes y la menor cantidad de huecos. Permite el suministro de una penalización por creación de hueco y una penalización por extensión de hueco en unidades de bases concordantes. GAP presenta un miembro de la familia de los mejores alineamientos.

Otro método para determinar la mejor concordancia global entre una secuencia consultada y una secuencia sujeto, también referido como un alineamiento de secuencias global, puede determinarse usando el programa informático FASTDB basado en el algoritmo de Brutlag et al. (*Comp. App. Biosci.* 6:237-245 (1990)). En un alineamiento de secuencias, las secuencias consultadas y sujeto son ambas secuencias de ADN. Una secuencia de ARN puede compararse convirtiendo los U en T. El resultado de dicho alineamiento de secuencias global es un porcentaje de identidad.

Los algoritmos BLAST y BLAST 2.0, pueden usarse para determinar el porcentaje de identidad de secuencia y similitud de secuencia. Éstos están descritos en Altschul et al. (1977) *Nuc. Acids Res.* 25:3389-3402, y Altschul et al (1990) *J. Mol. Biol.* 215:403-410, respectivamente. El software para llevar a cabo los análisis BLAST está disponible públicamente a través del National Center for Biotechnology Information. Este algoritmo implica en primer lugar identificar parejas de secuencias con alta puntuación (HSP) mediante la identificación de palabras cortas con longitud W en la secuencia consultada, que o bien concuerdan o satisfacen una puntuación de umbral con valor positivo T cuando se alinea con una palabra de la misma longitud en una secuencia de la base de datos. T se refiere como el umbral de puntuación de la palabra vecina (Altschul et al, supra). Estos logros de palabra vecina iniciales actúan como semillas para iniciar búsquedas para encontrar HSP más largas que las contienen. Los aciertos de palabra se extienden en ambas direcciones a lo largo de cada secuencia hasta que pueda incrementarse la puntuación acumulativa de alineamiento. Las puntuaciones acumulativas se calculan usando, para secuencias de nucleótidos, los parámetros M (puntuación de recompensa para una pareja de residuos concordantes; siempre > 0) y N (puntuación de penalización para residuos no concordantes; siempre < 0). Para secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabra en cada dirección se para cuando: la puntuación acumulativa de alineamiento cae por la cantidad X desde su máximo valor conseguido; la puntuación acumulativa va a cero o menos, debido a la acumulación de uno o más alineamientos de residuos con puntuación negativa; o se alcance el final de cualquiera de las secuencias. Los parámetros del algoritmo BLAST W, T, y X determinan la sensibilidad y velocidad del alineamiento. El programa BLASTN (para secuencias de nucleótidos) usa como defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, M=5, N=-4 y una comparación de ambas cadenas. Para secuencias de aminoácidos, el programa BLASTP usa como defecto una longitud de palabra de 3, y expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff (1989) *Proc. Natl. Acad. Sci USA* 89:10915) alineamientos (B) de 50, expectativa (E) de 10, M=5, N=-4, y una comparación de ambas cadenas. [0028] El algoritmo BLAST también lleva a cabo un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5787). Una medida de la similitud proporcionada por el algoritmo BLAST es la probabilidad menor de suma (P(N)), que proporciona una indicación de la probabilidad por la cual ocurriría al azar una concordancia entre dos secuencias de nucleótidos o aminoácidos. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad menor de suma en una comparación del ácido nucleico de ensayo con el ácido nucleico de referencia es menor de aproximadamente 0,2, más preferiblemente menor de aproximadamente 0,01, y lo más preferiblemente menor de aproximadamente 0,001.

La invención también contempla fragmentos de los polipéptidos descritos en la presente memoria. Un "fragmento" de polipéptido es una molécula de polipéptido que codifica un constituyente o es un constituyente de un polipéptido de la invención o variante de éste. Típicamente, el fragmento posee una actividad biológica cualitativa en común con el polipéptido del que es un constituyente. El fragmento de péptido puede tener una longitud entre aproximadamente 5 a aproximadamente 3.000 aminoácidos, entre aproximadamente 5 a aproximadamente 2.750 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 2.500 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 2.250 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 2.000 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 1.750 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 1.500 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 1.250 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 1.000 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 900 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 800 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 700 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 600 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 500 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 450 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 400 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 350 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 300 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 250 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 200 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 175 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 150 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 125 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 100 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 75 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 50 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 40 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 30 aminoácidos de longitud, entre aproximadamente 5 a aproximadamente 20 aminoácidos de longitud, y entre aproximadamente 5 a aproximadamente 15 aminoácidos de longitud. Alternativamente, el fragmento de péptido puede tener una longitud entre aproximadamente 5 a aproximadamente 10 aminoácidos.

También se contemplan fragmentos de los polinucleótidos descritos en la presente memoria. Un "fragmento" de polinucleótido es una molécula de polinucleótido que codifica un constituyente o es un constituyente de un polinucleótido de la invención o variante de éste. Los fragmentos de un polinucleótido no necesitan necesariamente codificar polipéptidos que retienen la actividad biológica. El fragmento, puede ser, por ejemplo, útil como una sonda de hibridación o cebador de PCR. El fragmento puede derivar de un polinucleótido de la invención o alternativamente puede sintetizarse por algún otro medio, por ejemplo por síntesis química.

Determinadas realizaciones de la invención se refieren a fragmentos de SEQ ID NO: 1. Un fragmento de SEQ ID NO: 1 puede comprender, por ejemplo, un constituyente de SEQ ID NO: 1 en el que el gen orf1 de la región del extremo 5' del gen está ausente. Alternativamente, un fragmento de SEQ ID NO: 1 puede comprender, por ejemplo, un constituyente de SEQ ID NO: 1 en el que el gen hisA de la región del extremo 3' del gen está ausente. Alternativamente, un fragmento de SEQ ID NO: 1 puede comprender, por ejemplo, un constituyente de SEQ ID NO: 1 en el que el gen orfA de la región del extremo 3' del gen está ausente. Alternativamente, un fragmento de SEQ ID NO: 1 puede comprender, por ejemplo, un constituyente de SEQ ID NO: 1 en el que el gen orf1 de la región del extremo 5' del gen está ausente y el gen orfA de la región del extremo 3' del gen está ausente. Alternativamente, un fragmento de SEQ ID NO: 1 puede comprender, por ejemplo, un constituyente de SEQ ID NO: 1 en el que el gen orf1 de la región del extremo 5' del gen está ausente y los genes hisA y/o orfA de la región del extremo 3' están ausentes.

En otras realizaciones, un fragmento de SEQ ID NO: 1 puede comprender uno o más marcos de lectura abiertos SXT. El marco de lectura abierto SXT puede seleccionarse del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes de éstos.

Las realizaciones adicionales de la invención se refieren a fragmentos de SEQ ID NO: 80. El fragmento de SEQ ID NO: 80 puede comprender uno o más marcos de lectura abiertos CYR. El marco de lectura abierto CYR puede seleccionarse del grupo que consiste en SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes de éstos.

En realizaciones particulares, los polinucleótidos de la invención pueden clonarse en un vector. El vector puede comprender, por ejemplo, una secuencia de ADN, ARN o ADN complementario (ADNc). El vector puede ser un vector plasmídico, un vector viral, o cualquier otro vehículo adecuado adaptado para la inserción de secuencias extrañas, su introducción en células y la expresión de las secuencias introducidas. Típicamente, el vector es un vector de expresión y puede incluir secuencias de control de la expresión y procesamiento tales como un promotor, un potenciador, sitios de unión a ribosomas, señales de poliadenilación y secuencias de terminación de la transcripción. La invención también contempla células huésped transformadas por dichos vectores. Por ejemplo, los

polinucleótidos de la invención pueden clonarse en un vector que se transforma en una célula huésped bacteriana, por ejemplo, E. coli. Los métodos para la construcción de vectores y su transformación en células huésped son conocidos generalmente en la técnica y se describen, por ejemplo, en Molecular Cloning: A Laboratory Manual (2a ed., Cold Spring Harbor Laboratory Press, Plainview, Nueva York, y, Ausubel F. M. et al. (Eds) Current Protocols in Molecular Biology (2007), John Wiley and Sons, Inc.

Sondas de nucleótidos, cebadores y anticuerpos

La invención contempla nucleótidos y fragmentos basados en las secuencias de los polinucleótidos descritos en la presente memoria para uso como cebadores y sondas para la identificación de secuencias homólogas.

Los nucleótidos y fragmentos pueden estar en la forma de oligonucleótidos. Los oligonucleótidos son cadenas cortas de residuos de nucleótidos adecuados para uso en reacciones de amplificación de ácidos nucleicos tales como PCR, teniendo típicamente una longitud de al menos aproximadamente 5 nucleótidos a aproximadamente 80 nucleótidos, más típicamente aproximadamente 10 nucleótidos de longitud a aproximadamente 50 nucleótidos de longitud, e incluso más típicamente aproximadamente 15 nucleótidos de longitud a aproximadamente 30 nucleótidos de longitud.

Las sondas son secuencias de nucleótidos con una longitud variable, por ejemplo entre aproximadamente 10 nucleótidos y varios miles de nucleótidos, para uso en la detección de secuencias homólogas, típicamente por hibridación. Las sondas de hibridación pueden ser fragmentos de ADN genómico, fragmentos de ADNc, fragmentos de ARN, u otros oligonucleótidos.

Los métodos para el diseño y/o producción de sondas de nucleótidos y/o cebadores son generalmente conocidos en la técnica, y se describen en Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2a ed., Cold Spring Harbor Laboratory Press, Plainview, Nueva York; Itakura K. et al. (1984) Annu. Rev. Biochem. 53:323; Innis et al., (Eds) (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, Nueva York); Innis y Gelfand, (Eds) (1995) PCR Strategies (Academic Press, Nueva York); e Innis y Gelfand, (Eds) (1999) PCR Methods Manual (Academic Press, Nueva York). Los cebadores y sondas de nucleótidos pueden prepararse, por ejemplo, por técnicas de síntesis química, por ejemplo, los métodos de fosfodiéster y fosfotriéster (véase, por ejemplo, Narang S. A. et al. (1979) Meth. Enzymol. 68:90; Brown, E. L. (1979) et al. Meth. Enzymol. 68:109; y la Patente U.S. No. 4356270), el método de dietilfosforamida (véase Beaucage S.L et al. (1981) Tetrahedron Letters, 22:1859-1862). Un método para sintetizar oligonucleótidos en un soporte sólido modificado se describe en la Patente U.S. No. 4458066.

Los ácidos nucleicos de la invención, incluyendo las sondas y cebadores mencionados anteriormente, pueden marcarse por la incorporación de un marcador para facilitar su detección. Las técnicas para marcar y detectar ácidos nucleicos se describen, por ejemplo, en Ausubel F. M. et al. (Eds) Current Protocols in Molecular Biology (2007), John Wiley and Sons, Inc. Los ejemplos de marcadores adecuados incluyen moléculas fluorescentes (por ejemplo, acetilaminofluoreno, 5-bromodesoxiuridina, digoxigenina, fluoresceína) e isótopos radiactivos (por ejemplo, ³²P, ³⁵S, ³H, ³³P). La detección del marcador puede conseguirse, por ejemplo, por técnicas químicas, fotoquímicas, inmunoquímicas, bioquímicas, o espectroscópicas.

Las sondas y cebadores de la invención pueden usarse, por ejemplo, para detectar o aislar cianobacterias y/o dinoflagelados en una muestra de interés. Adicionalmente o alternativamente, las sondas y cebadores de la invención pueden usarse para detectar o aislar un organismo cianotóxico y/o un organismo productor de cilindrospormisina en una muestra de interés. Adicionalmente o alternativamente, las sondas o cebadores de la invención pueden usarse para aislar secuencias correspondientes en otros organismos incluyendo, por ejemplo, otras especies bacterianas. Los métodos tales como la reacción en cadena de la polimerasa (PCR), hibridación, y semejantes pueden usarse para identificar dichas secuencias tomando como base su homología de secuencia con las secuencias mostradas en la presente memoria. Las secuencias que se seleccionan tomando como base su identidad de secuencia con las secuencias completas mostradas en la presente memoria o con fragmentos de éstas están englobadas por las realizaciones. Dichas secuencias incluyen secuencias que son ortólogos de las secuencias descritas. El término "ortólogos" se refiere a genes derivados de un gen ancestral común y que se encuentran en diferentes especies como resultado de la especiación. Los genes encontrados en diferentes especies se consideran ortólogos cuando sus secuencias de nucleótidos y/o sus secuencias de proteína codificadas comparten una identidad sustancial como se define en otro lugar de la presente memoria. Las funciones de los ortólogos están frecuentemente altamente conservadas entre las especies.

En técnicas de hibridación, toda o parte de una secuencia de nucleótidos conocida se usa para generar una sonda que hibrida selectivamente con otras secuencias de ácido nucleico correspondientes presentes en una muestra dada. Las sondas de hibridación pueden ser fragmentos de ADN genómico, fragmentos de ADNc, fragmentos de ARN, u otros oligonucleótidos, y pueden marcarse con un marcador detectable. Así, por ejemplo, las sondas para hibridación pueden prepararse marcando oligonucleótidos sintéticos basados en las secuencias de la invención.

El nivel de homología (identidad de secuencia) entre la sonda y la secuencia diana se determinará principalmente por la astringencia de las condiciones de hibridación. En particular, la secuencia de nucleótidos usada como una

sonda puede hibridar con un homólogo u otra variante de un polinucleótido descrito en la presente memoria en condiciones de astringencia baja, astringencia media o astringencia alta. Existen numerosas condiciones y factores, muy conocidos para los expertos en la técnica, que pueden emplearse para alterar la astringencia de la hibridación. Por ejemplo, la longitud y naturaleza (ADN, ARN, composición de bases) del ácido nucleico que va a hibridar con un ácido nucleico especificado; la concentración de sales y otros componentes, tales como la presencia o ausencia de formamida, sulfato de dextrano, polietilen glicol etc; y alterar la temperatura de las etapas de hibridación y/o lavado.

Típicamente, las condiciones de hibridación astringentes serán aquellas en las que la concentración de sal es menor de aproximadamente 1,5 M ión Na, típicamente aproximadamente 0,01 a 1,0 M concentración de ión Na (u otras sales) a pH 7,0 a 8,3 y la temperatura es al menos aproximadamente 30°C para sondas cortas (por ejemplo, 10 a 50 nucleótidos) y al menos aproximadamente 60°C para sondas largas (por ejemplo, mayores de 50 nucleótidos). Las condiciones astringentes también pueden conseguirse con la adición de agentes desestabilizadores tales como formamida. Las condiciones de astringencia baja ejemplares incluyen hibridación con una disolución tampón de 30% a 35% formamida, 1 M NaCl, 1% SDS (dodecil sulfato de sodio) a 37°C, y un lavado en 1X a 2X SSC (20X SSC = 3,0 M NaCl/0,3 M citrato de trisodio) a 50°C a 55°C. Las condiciones de astringencia moderadas ejemplares incluyen hibridación en 40% a 45% formamida, 1,0 M NaCl, 1% SDS a 37°C, y un lavado en 0,5X a 1X SSC a 55°C a 60°C. Las condiciones de astringencia alta ejemplares incluyen hibridación en 50% formamida, 1 M NaCl, 1% SDS a 37°C, y un lavado final en 0,1X SSC a 60°C a 65°C durante al menos aproximadamente 20 minutos. Opcionalmente, los tampones de lavado pueden comprender aproximadamente 0,1% a aproximadamente 1% SDS. La duración de la hibridación es generalmente menor de aproximadamente 24 horas, habitualmente aproximadamente 4 a aproximadamente 12 horas.

En una estrategia de PCR, los cebadores de oligonucleótidos pueden diseñarse para usarse en reacciones de PCR para amplificar secuencias de ADN correspondientes a partir de ADNc o ADN genómico extraído de cualquier organismo de interés. Los métodos para diseñar cebadores de PCR y clonación por PCR se conocen generalmente en la técnica y se describen en Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2a ed., Cold Spring Harbor Laboratory Press, Plainview, Nueva York); Ausubel F. M. et al. (Eds) Current Protocols in Molecular Biology (2007), John Wiley and Sons, Inc; Maniatis et al. Molecular Cloning (1982), 280-281; Innis et al. (Eds) (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, Nueva York); Innis y Gelfand, (Eds) (1995) PCR Strategies (Academic Press, Nueva York); e Innis y Gelfand, (Eds) (1999) PCR Methods Manual (Academic Press, Nueva York). Los métodos de PCR conocidos incluyen, pero no están limitados a, métodos usando cebadores emparejados, cebadores anidados, cebadores únicos específicos, cebadores degenerados, cebadores específicos de gen, cebadores específicos de vector, cebadores emparejados erróneamente parcialmente, y semejantes.

El experto en la técnica reconocerá que los cebadores descritos en la presente memoria para uso en PCR o RT-PCR también pueden usarse como sondas para la detección de secuencias SXT o CYR.

También se contemplan por la invención anticuerpos que son capaces de unirse específicamente a los polipéptidos de la invención. Los anticuerpos pueden usarse para detectar y analizar cualitativamente o cuantitativamente uno o más polipéptidos SXT o CYR en una muestra dada. Por "unirse específicamente" se entenderá que el anticuerpo es capaz de unirse al polipéptido diana o fragmento de éste con una afinidad mayor de la que se une a una proteína no relacionada. Por ejemplo, el anticuerpo puede unirse al polipéptido o fragmento de éste con una constante de unión en el intervalo de al menos aproximadamente 10^{-4} M a aproximadamente 10^{-10} M. Preferiblemente, la constante de unión es al menos aproximadamente 10^{-5} M, o al menos aproximadamente 10^{-6} M, más preferiblemente la constante de unión del anticuerpo para el polipéptido SXT o CYR o fragmento de éste es al menos aproximadamente 10^{-7} M, al menos aproximadamente 10^{-8} M, o al menos aproximadamente 10^{-9} M o más.

Los anticuerpos de la invención pueden existir en una variedad de formas, incluyendo por ejemplo como un anticuerpo completo, o como un fragmento de anticuerpo, u otro fragmento inmunológicamente activo de éste, tal como regiones determinantes de la complementariedad. De forma similar, el anticuerpo puede existir como un fragmento de anticuerpo que tiene dominios de unión a antígeno funcionales, esto es, dominios variables de cadena pesada y ligera. También, el fragmento de anticuerpo puede existir en una forma seleccionada del grupo que consiste en, pero no está limitada a: Fv, F_{ab}, F(ab)₂, scFv (Fv de cadena única), dAb (anticuerpo de dominio único), anticuerpos quiméricos, anticuerpos bi-específicos, fragmentos divalentes y trivalentes.

Un 'fragmento' de anticuerpo puede producirse por modificación de un anticuerpo completo o por síntesis del fragmento de anticuerpo deseado. Los métodos para generar anticuerpos, incluyendo fragmentos de anticuerpo, son conocidos en la técnica e incluyen, por ejemplo, síntesis por tecnología de ADN recombinante. El experto en la técnica tendrá conocimiento de los métodos para sintetizar anticuerpos, tales como los descritos, por ejemplo, en la Patente US No. 5296348 y Ausubel F. M. et al. (Eds) Current Protocols in Molecular Biology (2007), John Wiley and Sons, Inc.

Preferiblemente, los anticuerpos se preparan a partir de regiones o fragmentos discretos del polipéptido SXT o CYR de interés. Una parte antigénica de un polipéptido de interés puede tener cualquier longitud apropiada, tal como de aproximadamente 5 a aproximadamente 15 aminoácidos. Preferiblemente, una parte antigénica contiene al menos aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13 ó 14 residuos de aminoácidos.

En el contexto de esta especificación, la referencia a un anticuerpo específico para un polipéptido SXT o CYR de la invención incluye un anticuerpo que es específico para un fragmento del polipéptido de interés.

Los anticuerpos que se unen específicamente a un polipéptido de la invención pueden prepararse, por ejemplo, usando los polipéptidos SXT o CYR purificados o sus secuencias de ácido nucleico usando cualesquiera métodos adecuados conocidos en la técnica. Por ejemplo, un anticuerpo monoclonal, que contiene típicamente partes Fab, puede prepararse usando tecnología de hibridoma descrita en Harlow y Lane (Eds) *Antibodies - A Laboratory Manual*, (1988), Cold Spring Harbor Laboratory, N.Y.; Coligan, *Current Protocols in Immunology* (1991); Goding, *Monoclonal Antibodies: Principles and Practice* (1986) 2a ed; y Kohler y Milstein, (1975) *Nature* 256: 495-497. Dichas técnicas incluyen, pero no están limitadas a, la preparación de anticuerpos por la selección de anticuerpos de bibliotecas de anticuerpos recombinantes en fagos o vectores similares, así como la preparación de anticuerpos policlonales y monoclonales mediante la inmunización de conejos o ratones (véase, por ejemplo, Huse et al. (1989) *Science* 246: 1275-1281; Ward et al. (1989) *Nature* 341: 544-546).

También se entenderá que los anticuerpos de la invención incluyen anticuerpos humanizados, anticuerpos quiméricos y anticuerpos totalmente humanos. Un anticuerpo de la invención puede ser un anticuerpo bi-específico, que tiene especificidad de unión para más de un antígeno o epítipo. Por ejemplo, el anticuerpo puede tener especificidad para uno o más polipéptidos SXT o CYR o fragmentos de éstos, y además tiene especificidad de unión para otro antígeno. Los métodos para la preparación de anticuerpos humanizados, anticuerpos quiméricos, anticuerpos totalmente humanos, y anticuerpos biespecíficos son conocidos en la técnica e incluyen, por ejemplo, los descritos en la Patente de los Estados Unidos No. 6995243 presentada en 7 de febrero, 2006 de Garabedian, et al. y titulada "Antibodies that recognize and bind phosphorylated human glucocorticoid receptor and methods of using same".

Generalmente, una muestra que comprende potencialmente polipéptidos SXT o CYR puede ponerse en contacto con un anticuerpo que se une específicamente al polipéptido SXT o CYR o fragmento de éste. Opcionalmente, el anticuerpo puede fijarse a un soporte sólido para facilitar el lavado y aislamiento posterior del complejo, antes de poner en contacto el anticuerpo con una muestra. Los ejemplos de soportes sólidos incluyen, por ejemplo, placas de microtitulación, lechos, tics, o microlechos. Los anticuerpos también pueden unirse a una matriz ProteinChip o un sustrato sonda como se ha descrito anteriormente.

Los marcadores detectables para la identificación de anticuerpos unidos a los polipéptidos SXT o CYR de la invención incluyen, pero no están limitados a fluorocromos, agentes de tinción fluorescentes, radiomarcadores, enzimas tales como peroxidasa de rábano, fosfatasa alcalina y otros usados comúnmente en la técnica, y marcadores colorimétricos incluyendo oro coloidal o lechos de vidrio o plástico coloreados. Alternativamente, el marcador en la muestra puede detectarse usando un ensayo indirecto, en el que, por ejemplo, se usa un segundo anticuerpo marcado para detectar el anticuerpo unido específico del marcador.

Los métodos para detectar la presencia de o medir la cantidad de, un complejo anticuerpo-marcador incluyen, por ejemplo, la detección de fluorescencia, quimioluminiscencia, luminiscencia, absorbancia, birefringencia, transmitancia, reflectancia, o índice de refracción tal como resonancia de plasmón superficial, elipsometría, un método de espejo resonante, un método de guía de onda con acoplador de rejilla o interferometría. Los métodos de radio frecuencia incluyen espectroscopía de resonancia multipolar. Los métodos electroquímicos incluyen métodos de amperometría y voltametría. Los métodos ópticos incluyen métodos de formación de imágenes y métodos sin formación de imágenes y microscopía.

Los ensayos útiles para detectar la presencia de o medir la cantidad de, un complejo anticuerpo-marcador incluyen, por ejemplo, ensayo inmunoabsorbente ligado a enzima (ELISA), un ensayo radioinmune (RIA), o un ensayo de transferencia Western. Dichos métodos se describen, por ejemplo, en *Clinical Immunology* (Stites & Terr, eds., 7a ed. 1991); *Methods in Cell Biology: Antibodies in Cell Biology*, volumen 37 (Asai, ed. 1993); y Harlow y Lane, supra.

Métodos y kits para la detección

La invención proporciona métodos y kits para la detección y/o aislamiento de ácidos nucleicos y polipéptidos SXT. También se proporcionan métodos y kits para la detección y/o aislamiento de ácidos nucleicos y polipéptidos CYR.

En un aspecto, la invención proporciona un método para la detección de cianobacterias. El experto en la técnica entenderá que la detección de "cianobacterias" engloba la detección de una o más cianobacterias. El método comprende obtener una muestra para usarse en el método, y detectar la presencia de uno o más polinucleótidos o polipéptidos SXT como se describe en la presente memoria, o una variante o fragmento de éstos. La presencia de polinucleótidos, polipéptidos SXT, o variantes o fragmentos de éstos, es indicativa de cianobacterias en la muestra.

El polinucleótido SXT puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, ' SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ

ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas.

Alternativamente, el polinucleótido SXT puede ser un ARN o ADNc codificado por una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas y/o polipéptidos como se describe en la presente memoria, o una variante o fragmento de éstos.

El polipéptido SXT puede comprender una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 69, y variantes y fragmentos de éstas.

Los inventores han determinado que varios genes de la agrupación de genes SXT existen en organismos productores de saxitoxina y que están ausentes en los organismos con la agrupación de genes SXT que no producen saxitoxina. Específicamente, los inventores han identificado que el gen 6 (sxtA) (SEQ ID NO: 14), gen 9 (sxtG) (SEQ ID NO: 20), gen 10 (sxtH) (SEQ ID NO: 22), gen 11 (sxtI) (SEQ ID NO: 24) y gen 17 (sxtX) (SEQ ID NO: 36) de la agrupación de genes SXT están presentes sólo en organismos que producen saxitoxina.

De acuerdo con esto, en otro aspecto, la invención proporciona un método para detectar un organismo cianotóxico. El método comprende obtener una muestra para usarse en el método, y detectar un organismo cianotóxico tomando como base la detección de uno o más polinucleótidos SXT que comprenden una secuencia mostrada en SEQ ID NO: 14 (sxtA, gen 6), SEQ ID NO: 20 (sxtG, gen 9), SEQ ID NO: 22 (sxtH, gen 10), SEQ ID NO: 24 (sxtI, gen 11), SEQ ID NO: 36 (sxtX, gen 17), o variantes o fragmentos de éstos. Adicionalmente o alternativamente, un organismo cianotóxico puede detectarse tomando como base la detección de un ARN o ADNc que comprende una secuencia codificada por SEQ ID NO: 14 (sxtA, gen 6), SEQ ID NO: 20 (sxtG, gen 9), SEQ ID NO: 22 (sxtH, gen 10), SEQ ID NO: 24 (sxtI, gen 11), SEQ ID NO: 36 (sxtX, gen 17), o variantes o fragmentos de éstos. Adicionalmente o alternativamente, un organismo cianotóxico puede detectarse tomando como base la detección de uno o más polipéptidos que comprenden una secuencia mostrada en SEQ ID NO: 15 (SXTA), SEQ ID NO: 21 (SXTG), SEQ ID NO: 23 (SXTI), SEQ ID NO: 25 (SXTI), SEQ ID NO: 37 (SXTX), o variantes o fragmentos de éstas, en una muestra que se sospecha que comprende uno o más organismos cianotóxicos. El organismo cianotóxico puede ser cualquier organismo capaz de producir saxitoxina. En una realización preferida de la invención, el organismo cianotóxico es una cianobacteria o un dinoflagelado.

En determinadas realizaciones de la invención, los métodos para detectar cianobacterias o los métodos para detectar organismos cianotóxicos pueden comprender además la detección de uno o más polinucleótidos CYR o polipéptidos CYR como se describe en la presente memoria, o una variante o fragmento de éstos. El polinucleótido CYR puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes o fragmentos de éstas.

Alternativamente, el polinucleótido CYR puede ser un ARN o ADNc codificado por una secuencia de polinucleótido seleccionada del grupo que consiste en SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes o fragmentos de éstas.

El polipéptido CYR puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, y SEQ ID NO: 110, y variantes o fragmentos de éstas.

Los inventores han determinado que el gen 8 (cyrJ) (SEQ ID NO: 95) de la agrupación de genes CYR existe en organismos productores de cilindrospermopsina, y está ausente en organismos con la agrupación de genes CYR que no producen cilindrospermopsina. De acuerdo con esto, los métodos para detectar cianobacterias o los métodos para detectar organismos cianotóxicos pueden comprender además la detección de un organismo productor de cilindrospermopsina tomando como base la detección de un polinucleótido CYR que comprende una secuencia mostrada en SEQ ID NO: 95, o una variante o fragmento de éste. Adicionalmente o alternativamente, los métodos para detectar cianobacterias o los métodos para detectar organismos cianotóxicos pueden comprender además la detección de un organismo productor de cilindrospermopsina tomando como base la detección de un ARN o ADNc que comprende una secuencia codificada por SEQ ID NO: 95, o una variante o fragmento de éste. Adicionalmente o

alternativamente, los métodos para detectar cianobacterias o los métodos para detectar organismos cianotóxicos pueden comprender además la detección de un organismo productor de cilindrospermopsina tomando como base la detección de un polipéptido CYR que comprende una secuencia mostrada en SEQ ID NO: 96, o una variante o fragmento de éste.

5 En otro aspecto, la invención proporciona un método para la detección de cianobacterias. El experto en la técnica entenderá que la detección de "cianobacterias" engloba la detección de una o más cianobacterias. El método comprende obtener una muestra para usarse en el método, y detectar la presencia de uno o más polinucleótidos o polipéptidos CYR como se describe en la presente memoria, o una variante o fragmento de éstos. La presencia de polinucleótidos, polipéptidos CYR, o variantes o fragmentos de éstos, es indicativa de cianobacterias en la muestra.

10 El polinucleótido CYR puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes y fragmentos de éstas.

15 Alternativamente, el polinucleótido CYR puede ser un ARN o ADNc codificado por una secuencia seleccionada del grupo que consiste en SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes y fragmentos de éstas.

20 El polipéptido CYR puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, y SEQ ID NO: 110, y variantes o fragmentos de éstas.

25 En otro aspecto de la invención, se proporciona un método para detectar un organismo productor de cilindrospermopsina tomando como base la detección del gen 8 de CYR (cyrJ). El método comprende obtener una muestra para usarse en el método, y detectar la presencia de un polinucleótido CYR que comprende una secuencia mostrada en SEQ ID NO: 95, o una variante o fragmento de éste. Adicionalmente o alternativamente, el método para detectar un organismo productor de cilindrospermopsina tomando como base la detección del gen 8 de CYR (cyrJ) puede comprender la detección de un ARN o ADNc que comprende una secuencia codificada por SEQ ID NO: 95, o una variante o fragmento de éste. Adicionalmente o alternativamente, el método para detectar un organismo productor de cilindrospermopsina tomando como base la detección del gen 8 de CYR (cyrJ) puede comprender la detección de un polipéptido CYR que comprende una secuencia mostrada en SEQ ID NO: 96, o una variante o fragmento de éste.

35 En determinadas realizaciones de la invención, los métodos para detectar cianobacterias que comprenden la detección de secuencias CYR o variantes o fragmentos de éstas comprende además la detección de uno o más polinucleótidos SXT o polipéptidos SXT como se describe en la presente memoria, o una variante o fragmento de éstos.

40 El polinucleótido SXT puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas.

45 Alternativamente, el polinucleótido SXT puede ser un ARN o ADNc codificado por una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas y/o polipéptidos como se describe en la presente memoria, o una variante o fragmento de éstos.

50 El polipéptido SXT puede comprender una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 69, y variantes y fragmentos de éstas.

En otro aspecto, la invención proporciona un método para la detección de dinoflagelados. El experto en la técnica entenderá que la detección de "dinoflagelados" engloba la detección de uno o más dinoflagelados. El método

comprende obtener una muestra para usarse en el método, y detectar la presencia de uno o más polinucleótidos o polipéptidos SXT como se describe en la presente memoria, o una variante o fragmento de éstos. La presencia de polinucleótidos, polipéptidos SXT, o variantes o fragmentos de éstos, es indicativa de dinoflagelados en la muestra.

El polinucleótido SXT puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas.

Alternativamente, el polinucleótido SXT puede ser un ARN o ADNc codificado por una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas y/o polipéptidos como se describe en la presente memoria, o una variante o fragmento de éstos.

El polipéptido SXT puede comprender una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 69, y variantes y fragmentos de éstas.

Se puede sospechar que una muestra para usarse según los métodos descritos en la presente memoria comprende uno o más organismos cianotóxicos. Los organismos cianotóxicos pueden ser una o más cianobacterias y/o uno o más dinoflagelados. Adicionalmente o alternativamente, puede sospecharse que una muestra para usarse según los métodos descritos en la presente memoria comprende una o más cianobacterias y/o uno o más dinoflagelados. Una muestra para usarse según los métodos descritos en la presente memoria puede ser una muestra comparativa o control, por ejemplo, una muestra que comprende una concentración o densidad conocida de cianobacterias y/o dinoflagelados, o una muestra que comprende una o más especies o cepas conocidas de cianobacterias y/o dinoflagelados.

Una muestra para usarse según los métodos descritos en la presente memoria puede derivar de cualquier fuente. Por ejemplo, una muestra puede ser una muestra medioambiental. La muestra medioambiental puede derivar, por ejemplo, de agua salada, agua dulce o una floración de algas verdeazuladas. Alternativamente, la muestra puede derivar de una fuente de laboratorio, tal como un cultivo, o una fuente comercial.

Los expertos en la técnica apreciarán que los métodos y kits descritos en la presente memoria son adecuados generalmente para detectar cualesquiera organismos en los que están presentes las agrupaciones de genes SXT y/o CYR. Las cianobacterias adecuadas a las que se pueden aplicar los métodos de la invención pueden seleccionarse de los órdenes Oscillatoriales, Chroococcales, Nostocales y Stigonematales. Por ejemplo, las cianobacterias pueden seleccionarse de los géneros Anabaena, Nostoc, Microcystis, Planktothrix, Oscillatoria, Phormidium, y Nodularia. Por ejemplo, las cianobacterias pueden seleccionarse de las especies *Cylindrospermopsis raciborskii* T3, *Cylindrospermopsis raciborskii* AWT205, *Aphanizomenon ovalisporum*, *Aphanizomenon flos-aquae*, *Aphanizomenon* sp., *Umezakia natans*, *Raphidiopsis curvata*, *Anabaena bergii*, *Lyngbya wollei*, y *Anabaena circinalis*. Los ejemplos de dinoflagelados adecuados a los que se pueden aplicar los métodos y kits de la invención pueden seleccionarse de los géneros *Alexandrium*, *Pyrodinium* y *Gymnodinium*. Los métodos y kits de la invención también pueden emplearse para el descubrimiento de nuevas especies o géneros hepatotóxicos en colecciones de cultivos o de muestras medioambientales. Los métodos y kits de la invención también pueden emplearse para detectar cianotoxinas que se acumulan en otros animales, por ejemplo, peces y moluscos.

La detección de polinucleótidos y polipéptidos SXT y CYR descritos en la presente memoria puede llevarse a cabo usando cualquier método adecuado. Por ejemplo, los métodos para la detección de polinucleótidos y/o polipéptidos SXT y CYR descritos en la presente memoria pueden implicar el uso de un cebador, sonda o anticuerpo específico para uno o más polinucleótidos y polipéptidos SXT y CYR. Las técnicas y ensayos adecuados en los que el experto en la técnica puede utilizar un cebador, sonda o anticuerpo específico para uno o más polinucleótidos y polipéptidos SXT y CYR incluyen, por ejemplo, la reacción en cadena de la polimerasa (y variaciones relacionadas de esta técnica), ensayos basados en anticuerpos tales como ELISA y citometría de flujo, y microscopía fluorescente. Los métodos por los que pueden identificarse los polipéptidos SXT y CYR descritos en la presente memoria se conocen generalmente en la técnica, y se describen por ejemplo en Coligan J. E. et al. (Eds) *Current Protocols in Protein Science* (2007), John Wiley and Sons, Inc; Walker, J. M., (Ed) (1988) *New Protein Techniques: Methods in Molecular Biology*, Humana Press, Clifton, N.J; y Scopes, R. K. (1987) *Protein Purification: Principles and Practice*, 3a. Ed.,

Springer-Verlag, Nueva York, N.Y. Por ejemplo, los polipéptidos SXT y CYR descritos en la presente memoria pueden detectarse por transferencia western o análisis espectrofotométrico. Otros ejemplos de métodos adecuados para la detección de polipéptidos SXT y CYR se describen, por ejemplo, en la Patente US No. 4683195, Patente US No. 6228578, Patente US No. 7282355, Patente US No. 7348147 y publicación PCT No. WO/2007/056723.

- 5 En una realización preferida de la invención, la detección de los polinucleótidos y polipéptidos SXT y CYR se consigue por la amplificación de ADN de la muestra de interés por la reacción en cadena de la polimerasa, usando cebadores que hibridan específicamente con la secuencia SXT y/o CYR, o una variante o fragmento de éstas, y detectando la secuencia amplificada.

- 10 Los ácidos nucleicos y polipéptidos para análisis usando los métodos y kits descritos en la presente memoria pueden extraerse de organismos bien en cultivo mixto o como especies individuales o aislados de género. De acuerdo con esto, los organismos pueden cultivarse antes del aislamiento del ácido nucleico y/o polipéptido o alternativamente el ácido nucleico y/o polipéptidos pueden extraerse directamente de muestras medioambientales, tales como muestras de agua o floraciones de algas verdeazuladas.

- 15 Los métodos adecuados para la extracción y purificación de ácidos nucleicos para análisis usando los métodos y kits de la invención se conocen generalmente en la técnica y se describen, por ejemplo, en Ausubel F. M. et al. (Eds) Current Protocols in Molecular Biology (2007), John Wiley and Sons, Inc; Neilan (1995) Appl. Environ. Microbiol. 61:2286-2291; y Neilan et al. (2002) Astrobiol. 2:271-280. El experto en la técnica apreciará fácilmente que la invención no está limitada a los métodos específicos para el aislamiento de ácidos nucleicos descritos en esas publicaciones y que otros métodos adecuados están englobados por la invención. La invención puede llevarse a
20 cabo sin el aislamiento del ácido nucleico antes del análisis del ácido nucleico.

- Los métodos adecuados para la extracción y purificación de polipéptidos para los propósitos de la invención se conocen generalmente en la técnica y se describen, por ejemplo, en Coligan J. E. et al. (Eds) Current Protocols in Protein Science (2007), John Wiley and Sons, Inc; Walker, J. M., (Ed) (1988) New Protein Techniques: Methods in Molecular Biology, Humana Press, Clifton, N.J.; y Scopes, R. k. (1987) Protein Purification: Principles and Practice,
25 3a. Ed., Springer-Verlag, Nueva York, N.Y. Los ejemplos de técnicas adecuadas para la extracción de proteínas incluyen, pero no están limitadas a diálisis, ultrafiltración y precipitación. Las técnicas de purificación de proteínas adecuadas para usarse incluyen, pero no están limitadas a, cromatografía en fase reversa, cromatografía de interacción hidrofóbica, centrifugación, filtración en gel, precipitación con sulfato de amonio e intercambio iónico.

- Según los métodos y kits de la invención, los polinucleótidos SXT y CYR o variantes o fragmentos de éstos pueden detectarse por cualquier medio adecuado conocido en la técnica. En una realización preferida de la invención, los polinucleótidos SXT y CYR se detectan por amplificación por PCR. En la estrategia de PCR, pueden diseñarse
30 cebadores de oligonucleótidos para uso en las reacciones de PCR para amplificar los polinucleótidos SXT y CYR de la invención. También está englobada en la invención la amplificación por PCR de ADN complementario (ADNc) amplificado a partir de ARN mensajero (ARNm) derivado de la transcripción inversa de secuencias SXT y CYR (RT-PCR). Los métodos de PCR conocidos incluyen, pero no están limitados a, métodos usando cebadores emparejados, cebadores anidados, cebadores únicos específicos, cebadores degenerados, cebadores específicos de gen, cebadores específicos de vector, cebadores emparejados erróneamente parcialmente, y semejantes. Los métodos para diseñar cebadores de PCR y RT-PCR se conocen generalmente en la técnica y se describen, por ejemplo, en Ausubel F. M. et al. (Eds) Current Protocols in Molecular Biology (2007), John Wiley and Sons, Inc;
40 Maniatis et al. Molecular Cloning (1982), 280-281; Innis et al. (Eds) (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, Nueva York); Innis y Gelfand, (Eds) (1995) PCR Strategies (Academic Press, Nueva York); Innis y Gelfand, (Eds) (1999) PCR Methods Manual (Academic Press, Nueva York); y Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2a ed., Cold Spring Harbor Laboratory Press, Plainview, Nueva York).

- El experto en la técnica apreciará fácilmente que pueden alterarse diferentes parámetros de los procedimientos de PCR y RT-PCR sin afectar la capacidad para conseguir el producto deseado. Por ejemplo, puede variarse la concentración de sal o puede variarse el tiempo y/o temperatura de una o más de las etapas de desnaturalización, hibridación y extensión. De forma similar, la cantidad de molde de ADN, ADNc o ARN también puede variarse dependiendo de la cantidad de ácido nucleico disponible o la cantidad óptima de molde requerida para una amplificación eficaz. Los cebadores para uso en los métodos y kits de la presente invención son típicamente
50 oligonucleótidos, teniendo típicamente una longitud de al menos aproximadamente 5 nucleótidos a aproximadamente 80 nucleótidos, más típicamente aproximadamente 10 nucleótidos de longitud a aproximadamente 50 nucleótidos de longitud e incluso más típicamente aproximadamente 15 nucleótidos de longitud a aproximadamente 30 nucleótidos de longitud. El experto en la técnica reconocerá que los cebadores descritos en la presente memoria pueden ser útiles para varias aplicaciones diferentes, incluyendo pero no limitado a PCR, RT-PCR, y uso de sondas para la detección de secuencias SXT o CYR.
55

- Dichos cebadores pueden prepararse por cualquier método adecuado, incluyendo, por ejemplo, síntesis química directa o clonación y restricción de las secuencias apropiadas. No todas las bases en el cebador necesitan reflejar la secuencia de la molécula de molde con la que hibridará el cebador, el cebador sólo necesita contener suficientes bases complementarias para permitir que el cebador hibride con el molde. Un cebador también puede incluir bases con emparejamiento erróneo en una o más posiciones, siendo bases que no son complementarias a las bases en el
60

molde, sino que se diseñan para incorporar cambios en el ADN después de la extensión o amplificación de las bases. Un cebador puede incluir bases adicionales, por ejemplo, en la forma de una secuencia de reconocimiento de enzimas de restricción en el extremo 5', para facilitar la clonación del ADN amplificado.

- La invención proporciona un método para detectar un organismo cianotóxico tomando como base la detección de uno o más de SXT gen 6 (sxtA), SXT gen 9 (sxtG), SXT gen 10 (sxtH), SXT gen 11 (sxtI) y SXT gen 17 (sxtX) (SEQ ID NOS: 14, 20, 22, 24, y 36 respectivamente), o fragmentos o variantes de éstos. Adicionalmente o alternativamente, un organismo cianotóxico puede detectarse tomando como base la detección de uno o más de los siguientes polipéptidos SXT: SXTA (SEQ ID NO: 15), SXTG (SEQ ID NO: 21), SXTH (SEQ ID NO: 23), SXTI (SEQ ID NO: 25), SXTX (SEQ ID NO: 37), o fragmentos o variantes de éstos.
- El experto en la técnica reconocerá que cualesquiera cebadores capaces de amplificar las secuencias SXT y/o CYR indicadas, o variantes o fragmentos de éstas, son adecuados para uso en los métodos de la invención. Por ejemplo, las parejas de cebador de oligonucleótido adecuadas para la amplificación por PCR de SXT gen 6 (sxtA) puede comprender un primer cebador que comprende la secuencia de SEQ ID NO: 70 y un segundo cebador que comprende la secuencia de SEQ ID NO: 71, un primer cebador que comprende la secuencia de SEQ ID NO: 72 y un segundo cebador que comprende la secuencia de SEQ ID NO: 73, un primer cebador que comprende la secuencia de SEQ ID NO: 74 y un segundo cebador que comprende la secuencia de SEQ ID NO: 75, un primer cebador que comprende la secuencia de SEQ ID NO: 76 y un segundo cebador que comprende la secuencia de SEQ ID NO: 77, un primer cebador que comprende la secuencia de SEQ ID NO: 78 y un segundo cebador que comprende la secuencia de SEQ ID NO: 79, un primer cebador que comprende la secuencia de SEQ ID NO: 113 y un segundo cebador que comprende la secuencia de SEQ ID NO: 114, o un primer cebador que comprende la secuencia de SEQ ID NO: 115 o SEQ ID NO: 116 y un segundo cebador que comprende la secuencia de SEQ ID NO: 117.
- Las parejas de cebador de oligonucleótido adecuadas para la amplificación de SXT gen 9 (sxtG) pueden comprender un primer cebador que comprende la secuencia de SEQ ID NO: 118 y un segundo cebador que comprende la secuencia de SEQ ID NO: 119, o un primer cebador que comprende la secuencia de SEQ ID NO: 120 y un segundo cebador que comprende la secuencia de SEQ ID NO: 121.
- Las parejas de cebador de oligonucleótido adecuadas para la amplificación de SXT gen 10 (sxtH) pueden comprender un primer cebador que comprende la secuencia de SEQ ID NO: 122 y un segundo cebador que comprende la secuencia de SEQ ID NO: 123.
- Las parejas de cebador de oligonucleótido adecuadas para la amplificación de SXT gen 11 (sxtI) pueden comprender un primer cebador que comprende la secuencia de SEQ ID NO: 124 o SEQ ID NO: 125 y un segundo cebador que comprende la secuencia de SEQ ID NO: 126, o un primer cebador que comprende la secuencia de SEQ ID NO: 127 y un segundo cebador que comprende la secuencia de SEQ ID NO: 128.
- Las parejas de cebador de oligonucleótido adecuadas para la amplificación de SXT gen 17 (sxtX) pueden comprender un primer cebador que comprende la secuencia de SEQ ID NO: 129 y un segundo cebador que comprende la secuencia de SEQ ID NO: 130, o un primer cebador que comprende la secuencia de SEQ ID NO: 131 y un segundo cebador que comprende la secuencia de SEQ ID NO: 132.
- El experto en la técnica reconocerá qué fragmentos y variantes de las parejas de cebador mencionadas anteriormente también pueden amplificar eficazmente las secuencias de SXT gen 6 (sxtA), SXT gen 9 (sxtG), SXT gen 10 (sxtH), SXT gen 11 (sxtI) o SXT gen 17 (sxtX).
- En determinadas realizaciones de la invención, se detectan secuencias de polinucleótidos derivadas del gen de CYR tomando como base la detección de CYR gen 8 (cyrJ) (SEQ ID NO: 95). Las parejas de cebador de oligonucleótido adecuadas para la amplificación por PCR de CYR gen 8 (cyrJ) pueden comprender un primer cebador que tiene la secuencia de SEQ ID NO: 111 o un fragmento o variante de éste y un segundo cebador que tiene la secuencia de SEQ ID NO: 112 o un fragmento de éste.
- También están incluidas en el alcance de la presente invención variantes y fragmentos de los cebadores de oligonucleótido ejemplificados. El experto en la técnica también reconocerá que la invención no está limitada al uso de los cebadores específicos ejemplificados, y que también pueden usarse secuencias cebadoras alternativas, siempre que los cebadores se diseñen apropiadamente de manera que se permita la amplificación de secuencias SXT y/o CYR. Las secuencias de cebador adecuadas pueden determinarse por los expertos en la técnica usando procedimientos rutinarios sin experimentación excesiva. La localización de cebadores adecuados para la amplificación de secuencias SXT y/o CYR puede determinarse por factores tales como el contenido G+C y la capacidad de una secuencia de formar estructuras secundarias no deseadas.
- Los métodos adecuados de análisis de los ácidos nucleicos amplificados son muy conocidos para los expertos en la técnica y se describen, por ejemplo, en Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2a ed., Cold Spring Harbor Laboratory Press, Plainview, Nueva York); Ausubel F. M. et al. (Eds) *Current Protocols in Molecular Biology* (2007), John Wiley and Sons, Inc; y Maniatis et al. *Molecular Cloning* (1982), 280-281. Los métodos adecuados de análisis de los ácidos nucleicos amplificados incluyen, por ejemplo, electroforesis en gel que puede estar precedida o no de digestión con enzimas de restricción, y/o secuenciación de ácidos nucleicos. La

electroforesis en gel puede comprender electroforesis en gel de agarosa o electroforesis en gel de poliacrilamida, técnicas usadas comúnmente por los expertos en la técnica para la separación de fragmentos de ADN tomando como base el tamaño. La concentración de agarosa o poliacrilamida en el gel determina en gran medida la capacidad de resolución del gel y la concentración apropiada de agarosa o poliacrilamida dependerá por lo tanto del tamaño de los fragmentos de ADN que se quieren distinguir.

En otras realizaciones de la invención, los polinucleótidos SXT y CYR y variantes o fragmentos de éstos pueden detectarse por el uso de sondas adecuadas. Las sondas de la invención se basan en las secuencias de los polinucleótidos SXT y/o CYR descritas en la presente memoria. Las sondas son secuencias de nucleótidos con una longitud variable, por ejemplo entre aproximadamente 10 nucleótidos y varios miles de nucleótidos, para uso en la detección de secuencias homólogas, típicamente por hibridación. Las sondas de hibridación de la invención pueden ser fragmentos de ADN genómico, fragmentos de ADNc, fragmentos de ARN, u otros oligonucleótidos.

Los métodos para el diseño y/o producción de sondas de nucleótidos se conocen generalmente en la técnica, y se describen, por ejemplo, en Robinson P. J. et al. (Eds) *Current Protocols in Cytometry* (2007), John Wiley and Sons, Inc; Ausubel F. M. et al. (Eds) *Current Protocols in Molecular Biology* (2007), John Wiley and Sons, Inc; Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2a ed., Cold Spring Harbor Laboratory Press, Plainview, Nueva York; y Maniatis et al. *Molecular Cloning* (1982), 280-281. Las sondas de nucleótidos pueden prepararse, por ejemplo, por técnicas de síntesis química, por ejemplo, los métodos de fosfodiéster y fosfotriéster (véase, por ejemplo, Narang S. A. et al. (1979) *Meth. Enzymol.* 68:90; Brown, E. L. (1979) et al. *Meth. Enzymol.* 68:109; y la Patente U.S. No. 4356270), el método de dietilfosforamida (véase Beaucage S.L et al. (1981) *Tetrahedron Letters*, 22:1859-1862). Un método para sintetizar oligonucleótidos en un soporte sólido modificado se describe en la Patente U.S. No. 4458066.

Las sondas de la invención pueden marcarse por la incorporación de un marcador para facilitar su detección. Las técnicas para marcar y detectar ácidos nucleicos se describen, por ejemplo, en Ausubel F. M. et al. (Eds) *Current Protocols in Molecular Biology* (2007), John Wiley and Sons, Inc. Los ejemplos de marcadores adecuados incluyen moléculas fluorescentes (por ejemplo, acetilaminofluoreno, 5-bromodesoxiuridina, digoxigenina, fluoresceína) e isótopos radiactivos (por ejemplo, ³²P, ³⁵S, ³H, ³³P). La detección del marcador puede conseguirse, por ejemplo, por técnicas químicas, fotoquímicas, inmunoquímicas, bioquímicas, o espectroscópicas.

Los métodos y kits de la invención también engloban el uso de anticuerpos que son capaces de unirse específicamente a los polipéptidos de la invención. Los anticuerpos pueden usarse para detectar y analizar cualitativamente o cuantitativamente uno o más polipéptidos SXT o CYR en una muestra dada. Los métodos para la generación y uso de anticuerpos son conocidos generalmente en la técnica y se describen, por ejemplo, en Harlow y Lane (Eds) *Antibodies - A Laboratory Manual*, (1988), Cold Spring Harbor Laboratory, N.Y: Coligan, *Current Protocols in Immunology* (1991); Goding, *Monoclonal Antibodies: Principles and Practice* (1986) 2a ed; y Kohler y Milstein, (1975) *Nature* 256: 495-497. Los anticuerpos pueden conjugarse con un fluorocromo que permite la detección, por ejemplo, por citometría de flujo, inmunohistoquímica u otros medios conocidos en la técnica. Alternativamente, el anticuerpo puede unirse a un sustrato que permite la detección colorimétrica o quimioluminiscente. La invención también contempla el uso de anticuerpos secundarios capaces de unirse a uno o más anticuerpos capaces de unirse específicamente a los polipéptidos de la invención.

La invención también proporciona kits para la detección de organismos cianotóxicos y/o cianobacterias, y/o dinoflagelados. En general, los kits de la invención comprenden al menos un agente para detectar la presencia de uno más polinucleótidos o polipéptidos SXT y/o CYR descritos en la presente memoria, o una variante o fragmento de éstos. En el kit puede incluirse cualquier agente adecuado capaz de detectar secuencias SXT y/o CYR de la invención. Los ejemplos no limitativos incluyen cebadores, sondas y anticuerpos.

En un aspecto, la invención proporciona un kit para la detección de cianobacterias, comprendiendo el kit al menos un agente para detectar la presencia de uno o más polinucleótidos o polipéptidos SXT como se describe en la presente memoria, o una variante o fragmento de éstos.

El polinucleótido SXT puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas.

Alternativamente, el polinucleótido SXT puede ser un ARN o ADNc codificado por una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y

variantes y fragmentos de éstas y/o polipéptidos como se describe en la presente memoria, o una variante o fragmento de éstos.

El polipéptido SXT puede comprender una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 69, y variantes y fragmentos de éstas.

También se proporciona un kit para la detección de organismos cianotóxicos. El kit comprende al menos un agente para detectar la presencia de uno o más polinucleótidos o polipéptidos SXT como se describe en la presente memoria, o una variante o fragmento de éstos.

El polinucleótido SXT puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, y variantes y fragmentos de éstas.

Alternativamente, el polinucleótido SXT puede ser un ARN o ADNc codificado por una secuencia seleccionada del grupo que consiste en SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, y variantes y fragmentos de éstas.

El polipéptido SXT puede comprender una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 37, y variantes y fragmentos de éstas.

El al menos un agente puede ser cualquier reactivo adecuado para la detección de polinucleótidos y/o polipéptidos SXT descritos en la presente memoria. Por ejemplo, el agente puede ser un cebador, un anticuerpo o una sonda. Sólo como ejemplificación, los cebadores o sondas pueden comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, y variantes y fragmentos de éstas.

En determinadas realizaciones de la invención, los kits para la detección de cianobacterias u organismos cianotóxicos pueden comprender además al menos un agente adicional capaz de detectar una o más secuencias de polinucleótido CYR y/o polipéptido CYR como se describe en la presente memoria, o una variante o fragmento de éstas.

El polinucleótido CYR puede comprender un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes y fragmentos de éstas.

Alternativamente, el polinucleótido CYR puede comprender un ácido ribonucleico o ADN complementario codificado por una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes y fragmentos de éstas.

El polipéptido CYR puede comprender un polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, y SEQ ID NO: 110, y variantes y fragmentos de éstas.

El al menos un agente adicional puede seleccionarse, por ejemplo, del grupo que consiste en cebadores, anticuerpos y sondas. Un cebador o sonda adecuado puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 111, SEQ ID NO: 112, y variantes y fragmentos de éstas.

En otro aspecto, la invención proporciona un kit para la detección de cianobacterias, comprendiendo el kit al menos un agente para detectar la presencia de uno o más polinucleótidos o polipéptidos CYR como se describe en la presente memoria, o una variante o fragmento de éstos.

El polinucleótido CYR puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes y fragmentos de éstas.

Alternativamente, el polinucleótido CYR puede ser un ARN o ADNc codificado por una secuencia seleccionada del grupo que consiste en SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes y fragmentos de éstas.

- 5 El polipéptido CYR puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, y SEQ ID NO: 110, y variantes o fragmentos de éstas.

- 10 En determinadas realizaciones de la invención, los kits para detectar cianobacterias que comprenden uno o más agentes para la detección de secuencias CYR o variantes o fragmentos de éstas, pueden comprender además al menos un agente adicional capaz de detectar uno o más de los polinucleótidos SXT y/o polipéptidos SXT descritos en la presente memoria, o variantes o fragmentos de éstos.

- 15 El polinucleótido SXT puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas.

- 20 Alternativamente, el polinucleótido SXT puede ser un ARN o ADNc codificado por una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas y/o polipéptidos como se describe en la presente memoria, o una variante o fragmento de éstos.

- 25 El al menos un agente puede ser cualquier reactivo adecuado para la detección de polinucleótidos y/o polipéptidos CYR descritos en la presente memoria. Por ejemplo, el agente puede ser un cebador, un anticuerpo o una sonda. Sólo como ejemplificación, los cebadores o sondas pueden comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 111, SEQ ID NO: 112, y variantes y fragmentos de éstas.

También se proporciona un kit para la detección de dinoflagelados, comprendiendo el kit al menos un agente para detectar la presencia de uno o más polinucleótidos o polipéptidos SXT como se describe en la presente memoria, o una variante o fragmento de éstos.

- 35 El polinucleótido SXT puede comprender una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas.

- 45 Alternativamente, el polinucleótido SXT puede ser un ARN o ADNc codificado por una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, y variantes y fragmentos de éstas y/o polipéptidos como se describe en la presente memoria, o una variante o fragmento de éstos.

- 50 El polipéptido SXT puede comprender una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 69, y variantes y fragmentos de éstas.

En general, los kits de la invención pueden comprender cualquier número de componentes adicionales. Como ejemplos no limitativos, los componentes adicionales pueden incluir, reactivos para cultivo celular, muestras de referencia, tampones, marcadores, e instrucciones escritas para llevar a cabo el ensayo de detección.

Métodos de cribado

Los polipéptidos y polinucleótidos de la presente invención, y fragmentos y análogos de éstos son útiles para el cribado e identificación de compuestos y agentes que interaccionan con estas moléculas. En particular, los compuestos deseables son aquellos que modulan la actividad de estos polipéptidos y polinucleótidos. Dichos compuestos pueden ejercer un efecto modulador mediante la activación, estimulación, incremento, inhibición o prevención de la expresión o actividad de los polipéptidos y/o polinucleótidos. Los compuestos adecuados pueden ejercer su efecto mediante una interacción directa (por ejemplo, unión) o indirecta.

Los compuestos que se unen, o interaccionan de otra manera, con los polipéptidos y polinucleótidos de la invención, y específicamente los compuestos que modulan su actividad, pueden identificarse mediante una variedad de métodos adecuados. Los métodos no limitativos incluyen el método de dos híbridos, co-inmunoprecipitación, purificación por afinidad, espectroscopía de masas, purificación por afinidad en tándem, exposición en fago, transferencia de marcaje, micromatrices de ADN/coexpresión génica y micromatrices de proteínas.

Por ejemplo, un ensayo de dos híbridos puede usarse para determinar si un agente candidato o una pluralidad de agentes candidatos interacciona o se une a un polipéptido de la invención o una variante o fragmento de éste. El sistema de ensayo de dos híbridos en levaduras es un ensayo genético basado en levaduras usado típicamente para detectar interacciones proteína-proteína (Fields y Song., Nature 340: 245-246 (1989)). El ensayo usa la naturaleza multi-dominio de activadores transcripcionales. Por ejemplo, el dominio de unión a ADN de un activador transcripcional conocido puede fusionarse con un polipéptido de la invención o una variante o fragmento de éste, y el dominio de activación del activador transcripcional fusionarse con el agente candidato. La interacción entre el agente candidato y el polipéptido de la invención o una variante o fragmento de éste, acercará mucho los dominios de unión a ADN y de activación del activador transcripcional. La transcripción posterior de un gen informador específico activada por el activador transcripcional permite la detección de una interacción.

En una modificación de la técnica anterior, puede construirse una proteína de fusión fusionando el polipéptido de la invención o una variante o fragmento de éste a una etiqueta detectable, por ejemplo fosfatasa alcalina, y usar una forma modificada de inmunoprecipitación como describen Flanagan y Leder (Flanagan y Leder, Cell 63:185-194 (1990)).

Alternativamente, puede usarse co-inmunoprecipitación para determinar si un agente candidato o una pluralidad de agentes candidatos interacciona o se une a un polipéptido de la invención o una variante o fragmento de éste. Usando esta técnica, los organismos cianotóxicos, cianobacterias y/o dinoflagelados pueden lisarse en condiciones no desnaturalizantes adecuadas para la conservación de las interacciones proteína-proteína. La disolución resultante puede incubarse con un anticuerpo específico para un polipéptido de la invención o una variante o fragmento de éste e inmunoprecipitarse de la disolución en bruto, por ejemplo por captura con una proteína de unión a anticuerpo unida a un soporte sólido. La inmunoprecipitación del polipéptido de la invención o una variante o fragmento de éste por este método facilita la co-inmunoprecipitación de un agente asociado con esa proteína. La identificación de un agente asociado puede establecerse usando varios métodos conocidos en la técnica, incluyendo pero no limitado a SDS-PAGE, transferencia western y espectrometría de masas.

Alternativamente, puede usarse el método de exposición en fago para determinar si un agente candidato o una pluralidad de agentes candidatos interacciona o se une a un polipéptido de la invención o una variante o fragmento de éste. La exposición en fago es un ensayo para cribar interacciones proteicas mediante la integración de múltiples genes de un banco de genes en fagos. Con este método, se usan técnicas de ADN recombinante para expresar numerosos genes como fusiones con la proteína de cubierta de un bacteriófago de manera que el producto de péptido o proteína de cada gen se expone en la superficie de la partícula viral. De esta manera puede producirse una biblioteca completa de péptidos o productos proteicos de interés expuestos en fagos. Las bibliotecas resultantes de péptidos o productos proteicos expuestos en fagos puede cribarse para la capacidad de unirse a un polipéptido de la invención o una variante o fragmento de éste. El ADN extraído de fagos que interaccionan contiene las secuencias de las proteínas que interaccionan.

Alternativamente, puede usarse cromatografía de afinidad para determinar si un agente candidato o una pluralidad de agentes candidatos interacciona o se une a un polipéptido de la invención o una variante o fragmento de éste. Por ejemplo, un polipéptido de la invención o una variante o fragmento de éste, puede inmovilizarse en un soporte (tal como sefarosa) y pasar lisados celulares sobre la columna. Las proteínas que se unen al polipéptido inmovilizado de la invención o una variante o fragmento de éste pueden eluirse de la columna e identificarse, por ejemplo por secuenciación de aminoácidos N-terminal.

Los moduladores potenciales de la actividad de los polipéptidos de la invención pueden generarse por cribado por los métodos anteriores mediante varias técnicas conocidas para los expertos en la técnica. Por ejemplo, pueden usarse métodos tales como cristalografía con rayos X y espectroscopía de resonancia magnética nuclear para modelar la estructura de polipéptido de la invención o una variante o fragmento de éste, facilitando así el diseño de agentes moduladores potenciales usando modelado basado en ordenador. También pueden usarse varias formas de química combinatoria para generar posibles moduladores.

Los polipéptidos de la invención y las variantes o fragmentos de éstos apropiados, pueden usarse en cribados de alto rendimiento para ensayar compuestos candidatos para la capacidad de unirse a, o interaccionar de otra manera, con ellos. Estos compuestos candidatos pueden cribarse además frente a polipéptidos funcionales para determinar el efecto del compuesto sobre la actividad del polipéptido.

- 5 La presente invención también contempla compuestos que pueden ejercer su efecto modulador en polipéptidos de la invención alterando la expresión del polipéptido. En este caso, dichos compuestos pueden identificarse comparando el nivel de expresión del polipéptido en presencia de un compuesto candidato con el nivel de expresión en ausencia del compuesto candidato.

- 10 Se apreciará que los métodos descritos anteriormente son meramente ejemplos de los tipos de métodos que pueden utilizarse para identificar agentes capaces de interaccionar con, o modular la actividad de polipéptidos de la invención o variantes o fragmentos de éstos. Otros métodos adecuados serán conocidos para los expertos en la técnica y están en el alcance de esta invención.

- 15 Usando los métodos descritos anteriormente, puede identificarse que un agente es un agonista de un polipéptido de la invención o una variante o fragmento de éste. Los agentes que son agonistas aumentan una o más de las actividades biológicas del polipéptido. Alternativamente, los métodos descritos anteriormente pueden identificar un agente que es un antagonista de un polipéptido de la invención o una variante o fragmento de éste. Los agentes que son antagonistas retardan una o más de las actividades biológicas del polipéptido.

- 20 Los anticuerpos pueden actuar como agonistas o antagonistas de un polipéptido de la invención o una variante o fragmento de éste. Preferiblemente, los anticuerpos adecuados se preparan a partir de regiones o fragmentos discretos de los polipéptidos de la invención o variantes o fragmentos de éstos. Una parte antigénica de un polinucleótido de interés puede tener cualquier longitud apropiada, tal como aproximadamente 5 a aproximadamente 15 aminoácidos. Preferiblemente, una parte antigénica contiene al menos aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13 ó 14 residuos de aminoácidos.

- 25 Los métodos para la generación de anticuerpos adecuados serán apreciados fácilmente por los expertos en la técnica. Por ejemplo, puede prepararse un anticuerpo monoclonal específico para un polipéptido de la invención o una variante o fragmento de éste que contiene típicamente partes Fab, usando tecnología de hibridoma descrita en *Antibodies-A Laboratory Manual*, Harlow y Lane, eds., Cold Spring Harbor Laboratory, N.Y. (1988).

- 30 En esencia, en la preparación de anticuerpos monoclonales dirigidos frente a un polipéptido de la invención o una variante o fragmento de éste, puede usarse cualquier técnica que proporcione la producción de moléculas de anticuerpo por líneas celulares continuas en cultivo. Éstas incluyen la técnica de hibridoma desarrollada originalmente por Kohler et al., *Nature*, 256:495-497 (1975), así como la técnica de trioma, la técnica de hibridoma de células B humanas (Kozbor et al., *Immunology Today*, 4:72 (1983)), y la técnica de hibridoma EBV para producir anticuerpos monoclonales humanos (Cole et al., en *Monoclonal Antibodies and Cancer Therapy*, pp. 77- 96, Alan R. Liss, Inc., (1985)). Pueden crearse líneas celulares inmortales que producen anticuerpo por técnicas distintas de la fusión, tal como transformación directa de linfocitos B con ADN oncogénico, o transfección con el virus de Epstein-Barr. Véase, por ejemplo, M. Schreier et al., "Hybridoma Techniques" Cold Spring Harbor Laboratory, (1980); Hammerling et al., "Monoclonal Antibodies and T-cell Hybridomas" Elsevier/North-Holland Biochemical Press, Amsterdam (1981); y Kennett et al., "Monoclonal Antibodies", Plenum Press (1980).

- 40 Brevemente, un medio para producir un hibridoma a partir del cual se produce el anticuerpo monoclonal, un mieloma u otra línea celular que se perpetúe por sí misma se fusiona con linfocitos obtenidos del bazo de un mamífero hiperinmunizado con una parte de unión a un factor de reconocimiento de éste, o factor de reconocimiento, o una parte de unión a ADN específica de origen de éste. Los hibridomas que producen un anticuerpo monoclonal útil en la práctica de esta invención se identifican por su capacidad de inmunoreaccionar con los presentes factores de reconocimiento y su capacidad para inhibir actividad transcripcional especificada en células diana.

- 45 Un anticuerpo monoclonal útil en la práctica de la invención puede producirse iniciando un cultivo de hibridoma monoclonal que comprende un medio nutritivo que contiene un hibridoma que secreta moléculas de anticuerpo con la especificidad de antígeno apropiada. El cultivo se mantiene en condiciones y durante un periodo de tiempo suficiente para que el hibridoma secrete las moléculas de anticuerpo en el medio. El medio que contiene el anticuerpo se recoge. Las moléculas de anticuerpo pueden aislarse adicionalmente por técnicas muy conocidas.

- 50 De manera similar, existen varios procedimientos conocidos en la técnica que pueden usarse para la producción de anticuerpos policlonales. Para la producción de anticuerpos policlonales frente a un polipéptido de la invención o una variante o fragmento de éste, pueden inmunizarse varios animales huéspedes por inyección con un polipéptido de la invención, o una variante o fragmento de éste, incluyendo pero no limitado a conejos, pollos, ratones, ratas, ovejas, cabras, etc. Además, la variante de polipéptido o fragmento de éste puede conjugarse con un vehículo inmunogénico (por ejemplo, albúmina de suero bovino (BSA) o hemocianina de lapa (KLH)). También, pueden usarse varios adyuvantes para incrementar la respuesta inmunológica, incluyendo pero no limitado a adyuvante de Freund (completo e incompleto), geles minerales tales como hidróxido de aluminio, sustancias tensioactivas tales como risolecitina, polioles plurónicos, polianiones, péptidos, emulsiones de aceite, hemocianinas de lapa,
- 55

dinitrofenol, y adyuvantes humanos potencialmente útiles tales como BCG (bacilo Calmette-Guerin) y *Corynebacterium parvum*.

El cribado para el anticuerpo deseado también puede conseguirse por una variedad de técnicas conocidas en la técnica. Los ensayos para unión inmuno-específica de anticuerpos pueden incluir, pero no están limitados a, radioinmunoensayos, ELISA (ensayo inmunoabsorbente ligado a enzima), inmunoensayos en sandwich, ensayos inmunoradiométricos, reacciones de precipitación por difusión en gel, ensayos de inmunodifusión, inmunoensayos in situ, transferencias Western, reacciones de precipitación, ensayos de aglutinación, ensayos de fijación del complemento, ensayos de inmunofluorescencia, ensayos de proteína A, y ensayos de inmunoelectroforesis, y semejantes (véase, por ejemplo, Ausubel et al., *Current Protocols in Molecular Biology*, Vol. 1, John Wiley & Sons, Inc., Nueva York (1994)). La unión de anticuerpo puede detectarse mediante un marcador detectable en el anticuerpo primario. Alternativamente, el anticuerpo puede detectarse mediante su unión con un anticuerpo secundario o reactivo que está marcado apropiadamente. En la técnica se conoce una variedad de métodos para detectar la unión en un inmunoensayo y se incluyen en el alcance de la presente invención.

El anticuerpo (o fragmento de éste) producido frente a un polipéptido de la invención o una variante o fragmento de éste, tiene afinidad de unión para esa proteína. Preferiblemente, el anticuerpo (o fragmento de éste) tiene afinidad de unión o aidez mayor de aproximadamente $10^5 M^{-1}$, más preferiblemente mayor de aproximadamente $10^6 M^{-1}$, más preferiblemente aún mayor de aproximadamente $10^7 M^{-1}$ y lo más preferiblemente mayor de aproximadamente $10^8 M^{-1}$.

En términos de obtener una cantidad adecuada de un anticuerpo según la presente invención, se puede fabricar el o los anticuerpos usando fermentación discontinua con medio sin suero. Después de fermentar el anticuerpo puede purificarse mediante un procedimiento con múltiples etapas que incorpora etapas de cromatografía e inactivación/eliminación viral. Por ejemplo, el anticuerpo puede separarse en primer lugar por cromatografía por afinidad de Proteína A y tratarse con disolvente/detergente para inactivar cualesquiera virus con cubierta lipídica. Puede usarse purificación adicional, típicamente por cromatografía de intercambio aniónico o catiónico, para eliminar proteínas residuales, disolventes/detergentes y ácidos nucleicos. El anticuerpo purificado puede purificarse además y formularse en disolución salina al 0,9% usando columnas de filtración en gel. La preparación bruta formulada puede esterilizarse y filtrarse para virus y dispensarse.

Las realizaciones de la invención pueden utilizar tecnología antisentido para inhibir la expresión de un ácido nucleico de la invención o un fragmento o variante de éste mediante el bloqueo de la traducción del polipéptido codificado. La tecnología antisentido se aprovecha del hecho de que los ácidos nucleicos se emparejan con secuencias complementarias. Las moléculas antisentido adecuadas pueden fabricarse por síntesis química o, en el caso de ARN antisentido, por transcripción in vitro o in vivo cuando se une a un promotor, por métodos conocidos para los expertos en la técnica.

Por ejemplo, pueden generarse oligonucleótidos antisentido, típicamente con una longitud de 18-30 nucleótidos, que son al menos sustancialmente complementarios a lo largo de su longitud con una región de la secuencia de nucleótidos del polinucleótido de interés. La unión del oligonucleótido antisentido a sus secuencias de nucleótidos celulares complementarias puede interferir con la transcripción, procesamiento del ARN, transporte, traducción y/o estabilidad del ARNm. Pueden prepararse oligonucleótidos antisentido adecuados por métodos muy conocidos para los expertos en la técnica y pueden diseñarse para tomar como diana y unirse a regiones reguladoras de la secuencia de nucleótidos o a secuencias codificadoras (gen) o no codificadoras (región intergénica). Típicamente, los oligonucleótidos antisentido se sintetizarán en sintetizadores automáticos. Los oligonucleótidos antisentido adecuados pueden incluir modificaciones diseñadas para mejorar su administración a células, su estabilidad una vez dentro de una célula, y/o su unión a la diana apropiada. Por ejemplo, el oligonucleótido antisentido puede modificarse por la adición de una o más uniones fósforotioato, o la inclusión de uno o anillos morfolina en el núcleo (denominados oligonucleótidos 'morfolino').

Puede usarse una tecnología antisentido alternativa, conocida como ARN de interferencia (ARNi), según métodos conocidos en la técnica (véase, por ejemplo WO 99/49029 y WO 01/70949), para inhibir la expresión de un polinucleótido. ARNi se refiere a un medio de silenciamiento génico selectivo posterior a la transcripción por la destrucción de ARNm específicos por moléculas pequeñas de ARN de interferencia (ARNsi). El ARNsi se genera por la escisión de ARN bicatenario, en el que una cadena es idéntica al mensaje que se quiere inactivar. Pueden sintetizarse moléculas de ARN bicatenario en las que una cadena es idéntica a una región específica del transcrito de ARNm de p53 e introducirse directamente. Alternativamente, puede emplearse el ADNds correspondiente, que, una vez presentado intracelularmente se convierte en ARNds. Los métodos para la síntesis de moléculas adecuadas para uso en ARNi y para conseguir el silenciamiento génico posterior a la transcripción son conocidos para los expertos en la técnica.

Un medio adicional para inhibir la expresión puede conseguirse introduciendo construcciones de ácido nucleico antisentido catalítico, tales como ribozimas, que son capaces de escindir los transcritos de ARNm y evitar, de esta manera, la producción de proteína de tipo salvaje. Las ribozimas están dirigidas a e hibridan con una secuencia particular debido a dos regiones de complementariedad de secuencia con la diana que flanquea el sitio catalítico de la ribozima. Después de la unión, la ribozima escinde la diana de una manera específica de sitio. El diseño y ensayo

de ribozimas que reconocen y escinden específicamente secuencias de interés puede conseguirse por técnicas muy conocidas para los expertos en la técnica (véase, por ejemplo, Lieber y Strauss, 1995, *Molecular and Cellular Biology*, 15:540-551).

- 5 La invención se describirá ahora con referencia a ejemplos específicos, que no debe considerarse que limitan de ninguna manera el alcance de la invención.

Ejemplos

La invención se describirá ahora con referencia a ejemplos específicos, que no debe considerarse que limitan de ninguna manera el alcance de la invención.

Ejemplo 1: Cultivos de cianobacterias y caracterización de la agrupación de genes SXT.

- 10 Las cepas de cianobacterias usadas en el presente estudio (Figura 1) se crecieron en medio Jaworski en cultivo discontinuo estático a 26°C bajo iluminación continua ($10 \mu\text{moles m}^{-2} \text{s}^{-1}$). El ADN genómico total se extrajo de las células cianobacterianas por lisis con lisozima/SDS/proteinasa K seguido de extracción con fenol-cloroformo como se describe en Neilan, B. A. 1995. *Appl Environ Microbiol* 61:2286-2291. El ADN en el sobrenadante se precipitó en 2 volúmenes de etanol a - 20°C, se lavó con etanol al 70%, se disolvió en tampón TE (10:1), y se almacenó a - 20°C.
- 15 Las secuencias de cebadores de PCR usadas para la amplificación de ORF de *sxt* se muestran en la Figura 1B).

- Los amplicones de PCR se separaron por electroforesis en gel de agarosa en tampón TAE (40 mM Tris-acetato, 1 mM EDTA, pH 7,8), y se visualizaron por transiluminación UV después de teñir con bromuro de etidio (0,5 $\mu\text{g/ml}$). La secuenciación de regiones desconocidas de ADN se realizó por PCR mediada por adaptador como se describe en Moffitt et al. (2004) *Appl. Environ. Microbiol.* 70:6353-6362. La secuenciación automatizada del ADN se realizó
- 20 usando el sistema de secuenciación PRISM Big Dye cycle y un secuenciador modelo 373 (Applied Biosystems). Los datos de secuencia se analizaron usando software ABI Prism-Autoassembler, y se determinó el porcentaje de similitud e identidad con otras secuencias traducidas usando BLAST conjuntamente con National Center for Biotechnology Information (NIH), Fugue blast (<http://www-cryst.bioc.cam.ac.uk/fugue/>) se usó para identificar homólogos distantes mediante comparaciones secuencia-estructura. Las agrupaciones de genes *sxt* se
- 25 ensamblaron usando el software Phred, Phrap, y Consed (<http://www.phrap.org/phredphrapconsed.html>), y los marcos de lectura abiertos se identificaron manualmente. El número de registro GenBank para la agrupación de genes *sxt* de *C. raciborskii* T3 es DQ787200.

Ejemplo 2: Análisis espectrométrico de intermediarios SXT.

- Los extractos bacterianos y estándares SXT se analizaron por HPLC (Thermo Finnigan Surveyor HPLC y automuestreador) acoplada con un espectrómetro de masa con trampa de iones (Thermo Finnigan LCQ Deca XP Plus) ajustado con una fuente de electropulverización. La separación de los analitos se obtuvo en una columna 2,1 mm x 150 mm Phenomenex Luna 3 micrómetros C18 a 100 mL/min. El análisis se realizó usando un gradiente
- 30 empezando en 5% acetonitrilo en 10 mM ácido heptafluorobutírico (HFBA). Éste se mantuvo durante 10 min, y se subió hasta 100% acetonitrilo, durante 30 min. Las condiciones se mantuvieron a 100% acetonitrilo durante 10 min para lavar la columna y se volvió a 5% acetonitrilo en 10 mM HFBA y de nuevo se mantuvo durante 10 min para equilibrar la columna para la siguiente muestra. Esto resultó en un tiempo de corrida de 60 min por muestra. Se
- 35 inyectaron volúmenes de muestra de 10-100 mL para cada análisis. El eluato de HPLC se introdujo directamente en la fuente de electropulverización, que se programó como sigue: voltaje de electropulverización 5 kV, caudal de gas protector 30 unidades arbitrarias, caudal del gas auxiliar 5 unidades arbitrarias. La temperatura capilar fue 200°C y tenía un voltaje de 47 V. Las ópticas iónicas se optimizaron para una sensibilidad máxima antes del análisis de la muestra usando la función autoajuste del instrumento con una disolución de toxina estándar. Los espectros de masas se adquirieron en el modo centroide sobre el rango *m/z* 145-650. El ajuste del rango de masa fue 'normal', con 200 ms de tiempo de inyección máximo de ión y control de ganancia automático (AGC) encendido. Se
- 40 obtuvieron espectros de masa en tándem sobre un rango *m/z* relevante para el ión precursor. La energía de colisión fue típicamente 20-30 unidades arbitrarias ThermoFinnigan, y se optimizó para información máxima usando estándares cuando estaban disponibles.
- 45

Ejemplo 3: Identificación y secuenciación de la agrupación de genes SXT en *Cylindrospermopsis raciborskii* T3

- La O-carbamoyltransferasa se detectó inicialmente en *C. raciborskii* T3 mediante PCR degenerada, y posteriormente se denominó *sxtI*. La investigación adicional mostró que los homólogos de *sxtI* estaban presentes exclusivamente
- 50 en cepas productoras de toxina SXT de cuatro géneros de cianobacterias (Tabla 1), representando así un buen gen candidato en la biosíntesis de toxina SXT. La secuencia de la agrupación de genes biosintéticos de SXT (*sxt*) posible se obtuvo por paseo genómico en *sxtI* en *C. raciborskii* T3 (Figura 3). En *C. raciborskii* T3, esta agrupación de genes *sxt* abarca aproximadamente 35000 pb, que codifican 31 marcos de lectura abiertos (Figura 2). La agrupación también incluía otros genes que codifican enzimas de la biosíntesis de SXT, incluyendo una metiltransferasa (*sxtA1*), una aminotransferasa a clase II (*sxtA4*), una amidinotransferasa (*sxtG*), dioxigenasas (*sxtH*), además de la
- 55 O-carbamoyltransferasa (*sxtI*). El cribado por PCR de marcos de lectura abiertos de *sxt* seleccionados en cepas de cianobacterias tóxicas y no tóxicas mostró que estaban presentes exclusivamente en aislados productores de toxina SXT (Figura 1A), indicando la asociación de estos genes con el fenotipo tóxico. En los pasajes siguientes

describimos los marcos de lectura abiertos en la agrupación de genes *sxt* posible y sus funciones predichas, tomando como base análisis bioinformáticos, datos de LCMS/MS sobre intermedios biosintéticos y biosíntesis in vitro, cuando es aplicable.

Ejemplo 4: Predicción funcional de los genes biosintéticos de la molécula parental SXT

5 El análisis bioinformático de la agrupación de genes *sxt* reveló que contiene un ejemplo no descrito previamente de una estructura semejante a poliquétido sintasa (PKS), denominado *sxtA*. *SxtA* posee cuatro dominios catalíticos, *SxtA1* a *SxtA4*. Una búsqueda iterada en PSI-blast reveló una baja homología de secuencia de *SxtA1* con metiltransferasas dependientes de S-adenosilmetionina (SAM). Los análisis adicionales revelaron la presencia de tres restos de secuencia conservados en *SxtA1* (278-ITDMGCGDG- 286, 359-DPENILHI-366, y 424-VVNKHGLMIL-433) que son específicos para las metiltransferasas dependientes de SAM. *SxtA2* está relacionado con N-acetil transferasas relacionadas con GCN5 (GNAT). GNAT catalizan la transferencia de acetato de acetil-CoA a varios heteroátomos, y se ha indicado en asociación con otras PKS no convencionales, tales como *PedI*, en las que cargan la proteína transportadora de acilo (ACP) con acetato. *SxtA3* está relacionado con una ACP, y proporciona un sitio de unión a fosfopanteteinilo. *SxtA4* es homólogo a aminotransferasas de clase II y fue lo más similar a 8-amino-7-oxononanoato sintasa (AONS). Las aminotransferasas de clase II son un grupo monofilético de enzimas dependientes de piridoxal fosfato (PLP), y son las únicas enzimas que se sabe que realizan condensaciones de Claisen de aminoácidos. Por lo tanto, razonamos que *sxtA* realiza la primera etapa en la biosíntesis de SXT, que implica una condensación de Claisen.

20 La secuencia de reacción predicha de *SxtA*, tomando como base su estructura primaria, es la carga de la ACP (*SxtA3*) con acetato de acetil-CoA, seguido de la metilación catalizada por *SxtA1* de acetil-ACP, convirtiéndola en propionil-ACP. El dominio de aminotransferasa de clase II, *SxtA4*, realizaría entonces una condensación de Claisen entre propionil-ACP y arginina (Figura 4). El producto posible de *SxtA* es así 4-amino-3-oxoguanidinoheptano que se designa aquí como Compuesto A', (Figura 4). Para verificar esta ruta para la biosíntesis de SXT tomando como base análisis comparativos de secuencias génicas, se cribaron extractos celulares de *C. raciborskii* T3 por LC-MS/MS para la presencia del compuesto A' (Figura 5) así como arginina y SXT como controles. Se detectaron fácilmente arginina y SXT (Figura 5) y produjeron los iones de fragmento esperados. Por otra parte, los datos de LC-MS/MS obtenidos de m/z 187 fueron consistentes con la presencia de la estructura A de *C. raciborskii* T3 (Figura 5). Los espectros de MS/MS mostraron el ión fragmento esperado (m/z 170, m/z 128) después de la pérdida de amoniaco y guanidina de A'. Los datos de LC-MS/MS apoyaron fuertemente la función predicha de *SxtA* y así una reacción de iniciación revisada en la ruta de biosíntesis de SXT.

35 *sxtG* codifica una amidinotransferasa posible, que tenía la mayor similitud de secuencia de aminoácidos con L-arginina:lisina amidinotransferasas. Se propone que el producto de *SxtA* es el sustrato para la amidinotransferasa *SxtG*, que transfiere un grupo amidino de arginina al grupo α -amino A' (Figura 4), produciendo así 4,7-diguanidino-3-oxoheptano designado compuesto B' (Figura 3). Esta secuencia de reacciones hipotética también fue apoyada por la detección de C' por LC-MS/MS (Figura 4). Los extractos celulares de *C. raciborskii* T3, sin embargo, no contenían niveles mensurables de B' (4,7-diguanidino-3-oxoheptano). Una explicación posible para la falta de detección del intermedio B' es su ciclación rápida para formar C' a través de la acción de *SxtB*.

40 La agrupación de genes *sxt* codifica una enzima, *sxtB*, similar a las enzimas semejantes a citidina desaminasa de g-proteobacterias. El mecanismo catalítico de la citidina desaminasa es una escisión retro-aldol de amoniaco de citidina, que es el mismo mecanismo de reacción en la dirección inversa que la formación del primer heterociclo en la conversión de B' a C' (Figura 4). Por lo tanto, se sugiere que *SxtB* cataliza esta condensación semejante a retroaldol (etapa 4, Figura 4).

45 Se estudió la incorporación de metil metionina en SXT, y su hidroxilación. Sólo se retiene un hidrógeno derivado de metil metionina en SXT, y se observó un desplazamiento 1,2-H entre C-5 y C-6 derivados de acetato de SXT. La hidroxilación de la cadena lateral metilo del precursor de SXT se produce mediante la epoxidación de un doble enlace entre el grupo metilo derivado de SAM y el C-6 derivado de acetato. Este patrón de incorporación puede resultar de un ataque electrofílico de metil metionina en el doble enlace entre C-5 y C-6, que se habría formado durante la ciclación precedente. Posteriormente, la nueva cadena lateral metileno se epoxidaría, seguido de la apertura a un aldehído, y posterior reducción a un hidroxilo. La retención de un único hidrógeno derivado de metil metionina, el desplazamiento 1,2-H entre C-5 y C-6, y la ausencia de desplazamiento 1,2-H entre C-1 y C-5 es completamente consistente con los resultados de este estudio, en el que la introducción de metil metionina precede a la formación de los tres heterociclos.

55 *sxtD* codifica una enzima con similitud de secuencia con esteroide desaturasa y es la única desaturasa candidata presente en la agrupación de genes *sxt*, se predice que *SxtD* introduce un doble enlace entre C-1 y C-5 de C', y causa un desplazamiento 1,2-H entre C-5 y C-6 (compuesto D', Figura 3). El producto génico de *sxtS* tiene homología de secuencia con dioxigenasas dependientes de 2-oxoglutarato (2OG) con hierro no hemo. Éstas son enzimas multifuncionales que pueden realizar reacciones de hidroxilación, epoxidación, desaturación, ciclación, y expansión. Se ha indicado que las 2OG dioxigenasas catalizan la formación oxidativa de heterociclos. *SxtS* podría, por lo tanto, realizar la epoxidación consecutiva del nuevo doble enlace, y apertura del peróxido a un aldehído con biciclación simultánea. Esto explica la retención de un único hidrógeno derivado de metil metionina, y la ausencia de

un desplazamiento 1,2-H entre C-1 y C-5 de SXT (etapas 5 a 7, Figura 4). SxtU tiene similitud de secuencia con deshidrogenasas de alcohol de cadena corta. La enzima más similar con una función conocida es clavaldehído deshidrogenasa (AAF86624), que reduce el aldehído terminal de clavulanato-9-aldehído a un alcohol. Se predice, por lo tanto, que SxtU reduce el grupo aldehído terminal del precursor SXT en la etapa 8 (Figura 4), formando el compuesto E'.

La acción concertada de SxtD, SxtS y SxtU es, por lo tanto, la hidroxilación y biciclado del compuesto C' a E' (Figura 4). Apoyando esta ruta propuesta de la biosíntesis de SXT, LC-MS/MS obtenido de m/z 211 y m/z 225 permitieron la detección de los compuestos C' y E' de *C. raciborskii* T3 (Figura 5). Por otra parte, no se pudo encontrar evidencia por LC-MS/MS de los intermedios B (m/z 216), y C (m/z 198). Los espectros de MS/MS mostraron los iones fragmento esperados después de la pérdida de amoniaco y guanidina de C', así como la pérdida de agua en el caso de E'.

La detección de E' indicó que las reacciones finales que dan lugar a la molécula SXT completa son la O-carbamoiación de su grupo hidroxilo libre y una oxidación de C-12. La secuencia real de estas reacciones finales permanece, sin embargo, sin resolver. El producto génico de sxtI es lo más similar a una O-carbamoiiltransferasa predicha de *Trichodesmium erythraeum* (registro ABG50968) y otras O-carbamoiiltransferasas predichas de cianobacterias. Las O-carbamoiiltransferasas transfieren invariablemente un grupo carbamoilo de carbamoilfosfato a un grupo hidroxilo libre. Nuestros datos indican que SxtI puede catalizar la transferencia de un grupo carbamoilo de carbamoilfosfato al grupo hidroxilo libre de E'. No se encontraron homólogos de sxtJ y sxtK con una función conocida en las bases de datos, sin embargo, se indicó que los homólogos sxtJ y sxtK estaban codificados frecuentemente adyacentes a genes de O-carbamoiiltransferasa.

La agrupación de genes sxt contiene dos genes, sxtH y sxtT, cada uno codificando una subunidad de oxigenasa terminal de fenil-propionato bacteriano y dioxigenasas hidroxilantes de anillo relacionadas. El homólogo más cercano con una función predicha era capreomicidina hidroxilasa de *Streptomyces vinaceus*, que hidroliza un carbono del anillo (C-6) de capreomicidina. SxtH y SxtT pueden realizar, por lo tanto, una función similar en la biosíntesis de SXT, esto es, la oxidación o hidroxilación y oxidación de C-12, convirtiendo F' en SXT.

Los miembros que pertenecen a fenilpropionato bacteriano y dioxigenasas hidroxilantes de anillo relacionadas son enzimas multi-componente, ya que requieren una oxigenasa reductasa para su regeneración después de cada ciclo catalítico. La agrupación de genes sxt proporciona un sistema de transporte de electrones posible, que cumpliría con esta función. sxtV codifica una ferredoxina 4Fe-4S con alta homología de secuencia con una ferredoxina de *Nostoc punctiforme*. sxtW era lo más similar a las enzimas semejantes a fumarato reductasa/succinato deshidrogenasa de *A. variabilis* y *Nostoc punctiforme*, seguido de AsfA de *Pseudomonas putida*. AsfA y AsfB son enzimas implicadas en el transporte de electrones que resultan del catabolismo de aril sulfonatos. SxtV podría extraer posiblemente un par de electrones de succinato, convirtiéndolo en fumarato, y transferir los electrones a través de ferredoxina (SxtW) a SxtH y SxtT.

Ejemplo 5: Análisis comparativo de secuencia y asignación funcional de genes adaptados a SXT

Después de la síntesis de la molécula SXT parental, las enzimas de modificación introducen varios grupos funcionales. Además de SXT, *C. raciborskii* T3 produce toxinas N-1 hidroxiladas (neoSXT), descarbamoiadas (dcSXT), y N-sulfuriladas (GTX-5), mientras *A. circinalis* AWQC131C produce toxinas descarbamoiadas (dcSXT), O-sulfuriladas (GTX-3/2, dcGTX-3/2), así como tanto O- como N-sulfuriladas (C- 1/2), pero no toxinas N-1 hidroxiladas.

sxtX codifica una enzima con homología con cefalosporina hidroxilasa. sxtX sólo se detectó en *C. raciborskii* T3, *A. flos-aquae* NH- 5, y *Lyngbya wollei*, que producen análogos N-1 hidroxilados de SXT, tal como neoSXT. Este componente de la agrupación de genes no estaba presente en ninguna cepa de *A. circinalis*, y, por lo tanto, es probablemente la razón por la que esta especie no produce toxinas N-1 hidroxiladas PSP (Figura 1A). La función predicha de SxtX es, por lo tanto, la N-1 hidroxilación de SXT.

A. circinalis AWQC131C y *C. raciborskii* T3 también producen análogos N- y O-sulfatados de SXT (GTX-5, C-2/3, (dc)GTX- 3/4). Se ha descrito la actividad de dos sulfotransferasas dependientes de 3'-fosfato 5'-fosfosulfato (PAPS), que eran específicas para N- 21 de SXT y GTX-3/2, y O-22 de 11- hidroxil SXT, respectivamente, en el dinoflagelado productor de toxina SXT *Gymnodinium catenatum*. La agrupación de genes sxt de *C. raciborskii* T3 codifica una posible sulfotransferasa, SxtN. Una búsqueda PSI-BLAST con SxtN identificó sólo 25 proteínas hipotéticas de función desconocida con un valor E por encima del umbral (0,005). Una búsqueda de biblioteca de perfiles, sin embargo, reveló una relación estructural significativa de SxtN con estrógeno sulfotransferasa (1AQU) (puntuación Z=24,02) y otras sulfotransferasas. SxtN tiene una región N-terminal conservada, que corresponde a la región de unión de adenosina 3'-fosfato5'-fosfosulfato (PAPS) en 1AQU. No se sabe, sin embargo, si SxtN transfiere un grupo sulfato a N-21 o O-22. De forma interesante, la agrupación de genes sxt codifica una adenililsulfato quinasa (APSK), SxtO, cuyos homólogos están implicados en la formación de PAPS (Figura 2). APSK fosforila el producto de ATPsulfurilasa, adenililsulfato, convirtiéndolo en PAPS. Otras agrupaciones de genes biosintéticas que resultan en metabolitos secundarios sulfatados también contienen genes requeridos para la producción de PAPS.

Podrían producirse análogos descarbamoiados de SXT a través de uno de dos hipotéticos escenarios. Se propone que las enzimas que actúan aguas abajo de la carbamoiltransferasa, SxtI, en la biosíntesis de toxinas PSP tienen una amplia especificidad de sustrato, procesando precursores tanto carbamoilados como descarbamoiados de SXT. Alternativamente, puede producirse la escisión hidrolítica del resto carbamoilo de SXT o sus precursores. SxtL está relacionado con GDSL-lipasas, que son enzimas multifuncionales con actividades tioesterasa, arilesterasa, proteasa y lisofosfolipasa. La función de SxtL podría incluir, por lo tanto, la escisión hidrolítica del grupo carbamoilo de análogos de SXT.

Ejemplo 6: Genes SXT asociados en agrupaciones implicados en el transporte de metabolitos

sxtF y sxtM codificaban dos proteínas con alta similitud de secuencia con proteínas de extrusión multifármaco y de compuestos tóxicos (MATE) dirigidas por sodio de la familia NorM. Los miembros de la familia NorM de proteínas MATE son antiportadores dirigidos por sodio bacterianos, que exportan sustancias catiónicas. Todas las toxinas PSP son sustancias catiónicas, excepto las toxinas C que son zwitteriónicas. Por lo tanto, es probable que SxtF y SxtM también estén implicados en la exportación de toxinas PSP. Un estudio mutacional de NorM de *V. parahaemolyticus* identificó tres residuos cargados negativamente conservados (D32, E251, y D367) que confieren especificidad de sustrato, sin embargo, el mecanismo del reconocimiento del sustrato permanece desconocido. En SxtF, el residuo que corresponde a E251 de NorM está conservado, mientras los correspondientes a D32 y D367 están reemplazados por los aminoácidos neutros asparagina y tirosina, respectivamente. Los residuos correspondientes a D32 y E251 están conservados en SxtM, pero D367 está reemplazado por histidina. Los cambios en los residuos de unión a sustrato pueden reflejar las diferencias en los sustratos de la toxina PSP transportados por estas proteínas.

Ejemplo 7: Reguladores transcripcionales posibles de saxitoxina sintasa

Se ha indicado que los factores medioambientales, tales como disponibilidad de nitrógeno y fosfato regulan la producción de toxinas PSP en dinoflagelados y cianobacterias. Se identificaron dos factores transcripcionales, sxtY y sxtZ, relacionados con PhoU y OmpR, respectivamente, así como un regulador de dos componentes histidina quinasa próximos al extremo 3' de la agrupación de genes sxt en *C. raciborskii* T3. Las proteínas relacionadas con PhoU son reguladores negativos de la captación de fosfato mientras las proteínas semejantes a OmpR están implicadas en la regulación de una variedad de metabolismos, incluyendo nitrógeno y equilibrio osmótico. Por lo tanto, es probable que la producción de toxina PSP en *C. raciborskii* T3 esté regulada a nivel transcripcional en respuesta a la disponibilidad de fosfato, así como, otros factores medioambientales.

Ejemplo 8: Orígenes filogenéticos de los genes SXT

La agrupación de genes sxt de *C. raciborskii* T3 tiene una estructura de mosaico verdadera. Aproximadamente la mitad de los genes sxt de *C. raciborskii* T3 eran lo más similares a equivalentes de otras cianobacterias, sin embargo, los genes restantes tenían sus concordancias más cercanas con homólogos de proteobacterias, actinomicetes, esfingobacterias, y firmicutes. Existe un número de evidencias creciente de que la transferencia horizontal de genes (HGT) es una fuerza directriz principal detrás de la evolución de genomas procariotas, y se sabe que los genomas de cianobacterias están muy afectados por HGT, implicando frecuentemente transposasas y fagos. El hecho de que la mayoría de los genes sxt esté lo más relacionado con homólogos de otras cianobacterias, sugiere que la biosíntesis de SXT puede haber evolucionado en una cianobacteria ancestral que adquirió sucesivamente los genes restantes de otras bacterias a través de HGT. La organización estructural de la agrupación de genes sxt investigada, así como la presencia de varias transposasas relacionadas con la familia IS4, sugiere que casetes pequeños de genes sxt son móviles.

Ejemplo 9: Cultivos de cianobacterias y caracterización de la agrupación de genes CYR.

Las cepas de cianobacterias se crecieron en medio Jaworski como se describe en el Ejemplo 1 anterior. El ADN genómico total se extrajo de las células de cianobacterias por lisis con lisozima/SDS/proteinasa K seguido de extracción con fenol-cloroformo como se ha descrito previamente Neilan B. A. 1995.. Appl Environ Microbiol 61:2286-2291. El ADN en el sobrenadante se precipitó en 2 volúmenes de etanol a -20°C, se lavó con etanol al 70%, se disolvió en tampón TE (10:1), y se almacenó a -20°C.

La caracterización de regiones desconocidas de ADN que flanquean los genes posibles de biosíntesis de cilindrospermopsina se realizó usando una PCR mediada por adaptador como se describe en Moffitt et al. (2004) Appl. Environ. Microbiol. 70:6353-6362. Las PCR se realizaron en volúmenes de reacción de 20 µl que contenían 1 x tampón de Taq polimerasa 2,5 mM MgCl₂, 0,2 mM desoxinucleótidos trifosfato, 10 pmoles cada uno de los cebadores directo e inverso, entre 10 y 100 ng de ADN genómico y 0,2 U de Taq polimerasa (Fischer Biotech, Australia). El ciclado térmico se realizó en un ciclador térmico GeneAmp PCR System 2400 (Perkin Elmer Corporation, Norwalk, CT). El ciclado empezó con una etapa de desnaturalización a 94°C durante 3 min seguido de 30 ciclos de desnaturalización a 94°C durante 10 s, hibridación de cebador entre 55° y 65°C durante 20 s y una extensión de la cadena de ADN a 72°C durante 1-3 min. La amplificación se completó por una etapa final de extensión a 72°C durante 7 min. El ADN amplificado se separó por electroforesis en gel de agarosa en tampón TAE

(40 mM Tris-acetato, 1 mM EDTA, pH 7,8), y se visualizó por transiluminación UV después de tefir con bromuro de propidio (0,5 µg/ml).

La secuenciación automatizada del ADN se realizó usando el sistema de secuenciación PRISM Big Dye cycle y un secuenciador modelo 373 (Applied Biosystems, Foster City, CA). Los datos de secuencia se analizaron usando software ABI Prism-Autoassembler, mientras los valores de identidad/similitud con otras secuencias traducidas se determinaron usando BLAST conjuntamente con el National Center for Biotechnology Information (NIH, Bethesda, MD). Se usó Fugue blast (<http://www-cryst.bioc.cam.ac.uk/fugue/>) para identificar homólogos distantes a través de comparaciones secuencia-estructura. Las agrupaciones de genes se ensamblaron usando el software Phred, Phrap, y Consed (<http://www.phrap.org/phredphrapconsed.html>), y los marcos de lectura abiertos se identificaron manualmente. Los dominios poliquétido sintasa y péptido sintetasa no-ribosomal se determinaron usando las bases de datos especializadas en estructuras cristalinas (<http://www-ab.informatik.uni-tuebingen.de/software/NRPSpredictor>; <http://www.tigr.org/jravel/nrps/>, <http://www.nii.res.in/nrps-pks.html>).

Ejemplo 10: Cribado genético de cepas de cianobacterias productoras y no productoras de Cilindrospermopsina

Las cepas de cianobacterias productoras y no productoras de Cilindrospermopsina se cribaron para la presencia del gen de sulfotransferasa *cyrJ* usando el conjunto de cebadores *cynsulfF* (5' ACTTCTCTCCTTCCCTATC 3') (SEQ ID NO: 111) y *cynamR* (5' GAGTGAAAATGCGTAGAACTTG 3') (SEQ ID NO: 112). El ADN genómico se ensayó para amplificación positiva usando los cebadores del gen de 16S ARNr 27F y 809 como se describe en Neilan et al. (1997) Int. J. Syst. Bacteriol. 47:693-697. Los amplicones se secuenciaron, como se describe en el Ejemplo 9 anterior, para verificar la identidad del fragmento génico.

La biosíntesis de cilindrospermopsina implica una amidinotransferasa, un NRPS, y un PKS (AoaA, AoaB y AoaC, respectivamente). Con el fin de obtener la secuencia completa de la agrupación de genes de la biosíntesis de cilindrospermopsina, usamos tecnología de 'paseo genético' mediada por adaptador, iniciando el proceso a partir de una secuencia parcial del gen de amidinotransferasa de *C. raciborskii* AWT205. Se diseñaron cebadores externos sucesivos y se secuenció la agrupación de genes completa que abarca 43 kb, junto con 3,5 kb más a cada lado de la agrupación de genes de la toxina.

Estas regiones flanqueantes codifican genes auxiliares posibles (genes *hyp*), que incluyen chaperonas moleculares implicadas en la maduración de hidrogenasas. Debido al hecho de que estos genes están flanqueando la agrupación de genes de la cilindrospermopsina en ambos extremos, postulamos que la agrupación de genes de la toxina estaba insertada en esta área del genoma interrumpiendo así la agrupación de genes HYP. Esta reorganización genética está apoyada mecanísticamente por la presencia de secuencias semejantes a transposasa en la agrupación de cilindrospermopsina.

Se realizó el análisis informático de la agrupación de genes de la toxina y tomando como base la inferencia de la función génica usando alineamientos de secuencias (NCBI BLAST), homólogos estructurales predichos (Fugue Blast), y análisis de dominios PKS y NRPS usando servidores blast especializados basados en estructuras cristalinas. La agrupación de genes de la biosíntesis de cilindrospermopsina contiene 15 ORF, que codifican todas las funciones requeridas para la biosíntesis, regulación y exportación de la toxina cilindrospermopsina (Figura 6).

Ejemplo 11: Formación del esqueleto de carbono de CYR

La primera etapa en la formación del esqueleto de carbono de cilindrospermopsina implica la síntesis de guanidinoacetato a través de la transamidinación de glicina. *CyrA*, el homólogo AoaA, que codifica una amidinotransferasa similar a la arginina:glicina amidinotransferasa GATM humana, transfiere un grupo guanidino de una molécula donante, lo más probablemente arginina, a una molécula aceptora de glicina formando así guanidinoacetato (Figura 8, etapa 1).

La siguiente etapa (Figura 8, etapa 2) en la biosíntesis se realiza por *CyrB* (homólogo de AoaB), un NRPS-PKS mixto. *CyrB* abarca 8,7 kb y codifica los dominios siguientes; dominio de adenilación (dominio A) y una proteína transportadora de peptidilo (PCP) de un NRPS seguido de un dominio βquetosintasa (KS), dominio aciltransferasa (AT), dominio deshidratasa (DH), dominio metiltransferasa (MT), dominio quetoreductasa (KR), y una proteína transportadora de acilo (ACP) de origen PKS. *CyrB* por lo tanto debe catalizar la segunda reacción ya que es el único gen que contiene un dominio A que podría reclutar una unidad de inicio para extensiones PKS posteriores. El aminoácido específico activado por el dominio *CyrB* A no puede predecirse ya que sus residuos que confieren especificidad de sustrato no concuerdan con ninguna de las bases de datos disponibles (<http://www-ab.informatik.uni-tuebingen.de/software/NRPSpredictor>; <http://www.tigr.org/jravel/nrps/>, <http://www.nii.res.in/nrps-pks.html>). Hasta ahora, no se ha descrito ningún otro NRPS que utilice guanidinoacetato como sustrato. Se piensa que el dominio A activa el guanidinoacetato, que entonces se transfiere a través de un brazo oscilante de la proteína transportadora de peptidilo (PCP) al dominio KS. El dominio AT activa malonil-CoA y la une al ACP. Esto es seguido de una reacción de condensación entre el guanidinoacetato activado y malonil-CoA en el dominio KS. *CyrB* contiene dos moléculas reductoras, KR y DH. Su reacción concertada reduce el grupo ceto a un hidroxilo seguido de la eliminación de H₂O, lo que resulta en un enlace doble entre C13 y C14. El dominio metil transferasa (MT) identificado en *CyrB* a través de bases de datos NRPS/PKS (Ejemplo 9 anterior), es homólogo a MT dependiente de

S-adenosilmetionina (SAM). Por lo tanto, se sugiere que MT metila C13. Se propone que se produce un ataque nucleofílico del grupo amidino en N19 al nuevo enlace doble formado entre C13 y C14 a través de una 'adición de Michael'. La ciclación sigue las reglas de Baldwin para el cierre del anillo (Baldwin et al. (1997) J. Org. Chem 42:3846-3852), lo que resulta en la formación del primer anillo en cilindrospermopsina. Esta reacción podría ser espontánea y puede no requerir catálisis enzimática, ya que es energéticamente favorable. Ésta es la primera de tres formaciones de anillo.

La tercera etapa (Figura 8, etapa 3) en la biosíntesis implica CyrC (homólogo de AoaC), que codifica un PKS con dominios KS, AT, KR, y ACP. La acción de estos dominios resulta en la elongación de la cadena creciente por un acetato a través de la activación de malonil-CoA por el dominio AT, su transferencia a ACP y condensación en el dominio KS con el producto de CyrB. La cadena elongada se une al ACP de CyrC y el dominio KR reduce el grupo ceto a un grupo hidroxilo en C12. El módulo PKS que realiza esta etapa contiene un dominio KR y no contiene un dominio DH, esto sólo corresponde a CyrC.

Después de la catálisis de la enzima CyrC está CyrD (Figura 8, etapa 4), un PKS con cinco módulos; KS, AT, DH, KR, y un ACP. La acción de este módulo PKS en el producto de CyrC resulta en la adición de un acetato y la reducción del grupo ceto en C10 a un hidroxilo y deshidratación a un doble enlace entre C9 y C10. Este doble enlace es el sitio de un ataque nucleofílico por el grupo amidino N19 a través de otra adición de Michael que de nuevo sigue las reglas de Baldwin de cierre del anillo, lo que resulta en la formación del segundo anillo, el primer anillo de 6 miembros preparado en la cilindrospermopsina.

El producto de CyrD es el sustrato para CyrE (etapa 5 en la Figura 8), un PKS que contiene dominios KS, AT, DH, KR y un ACP. Como esta secuencia de dominios es idéntica a la de CyrD, no es posible en este estadio discernir qué PKS actúa en primer lugar, pero como se propone que su acción es idéntica es irrelevante en este punto. CyrE cataliza la adición de un acetato y la formación de un doble enlace entre C7 y C8. Este doble enlace es atacado por N18 a través de una adición de Michael y se produce la tercera ciclación, lo que resulta en el segundo anillo de 6 miembros.

CyrF es el módulo PKS final (etapa 6 de la Figura 8) y es un PKS mínimo que contiene sólo KS, AT, y ACP. CyrF actúa en el producto de CyrE y elonga la cadena por un acetato, dejando C4 y C6 sin reducir.

La etapa 7 en la ruta (Figura 8) implica la formación del anillo uracilo, una reacción que se requiere para la toxicidad del compuesto cilindrospermopsina final. La agrupación de genes de la cilindrospermopsina codifica dos enzimas con alta similitud de secuencia (87%) que se han denominado CyrG y CyrH. Una búsqueda Psi-blast (NCBI) seguida de una búsqueda en la biblioteca de perfiles Fugue (véase materiales y métodos) reveló que CyrG y CyrH son lo más similares a la familia de enzimas de amidohidrolasas/ureasas/dihidroorotasas, cuyos miembros catalizan la formación y escisión de enlaces N-C. Se propone que estas enzimas transfieren un segundo grupo guanidino de una molécula donante, tal como arginina o urea, a C6 y C4 de cilindrospermopsina lo que resulta en la formación del anillo uracilo. Estas enzimas realizan dos o tres reacciones dependiendo del donante de guanidino. La primera reacción consiste en la formación de un enlace covalente entre el N del donante de guanidino y C6 de cilindrospermopsina seguido de una eliminación de H₂O formando un doble enlace entre C5 y C6. La segunda reacción cataliza la formación de un enlace entre el segundo N en el donante de guanidino y C4 de cilindrospermopsina, simultáneamente con la rotura del enlace tioéster entre la proteína transportadora de acilo de CyrE y cilindrospermopsina, causando la liberación de la molécula del complejo enzimático. Los experimentos de siembra con acetato marcado han mostrado que el oxígeno en C4 tiene origen acetato y no se pierde durante la biosíntesis, por lo tanto requiere la formación de novo del anillo uracilo. La tercera reacción - si se requiere - catalizaría la escisión del grupo guanidino de una molécula donante distinta de urea. La acción de CyrG y CyrH en la formación del anillo uracilo en cilindrospermopsina describe una ruta de biosíntesis nueva de una pirimidina.

Una teoría sugiere un poliquétido lineal que asume fácilmente una conformación favorable para la formación de los anillos. La ciclación puede ser así espontánea y no está bajo control enzimático. Estos análisis muestran que esto puede ocurrir por etapas, con la formación sucesiva de anillos del intermedio apropiado al sintetizarse. Este mecanismo también explica la ausencia de un dominio tioesterasa o de ciclación, que están asociados habitualmente con módulos NRPS/PKS y catalizan la liberación y ciclación del producto final del complejo enzimático.

Ejemplo 12: Reacciones adaptadas a CYR

La biosíntesis de cilindrospermopsina requiere la acción de enzimas adaptadas con el fin de completar la biosíntesis, catalizando la sulfatación en C12 e hidroxilación en C7. El análisis de la agrupación de genes de cilindrospermopsina reveló tres enzimas candidatas para las reacciones adaptadas implicadas en la biosíntesis de cilindrospermopsina, concretamente CyrI, CyrJ, y CyrN. La sulfatación de cilindrospermopsina en C12 se realiza probablemente por la acción de una sulfotransferasa. CyrJ codifica una proteína que es lo más similar a sulfotransferasas dependientes de 3'-fosfoadenilil sulfato (PAPS) humanas. La agrupación de genes de cilindrospermopsina también codifica una adenilsulfato quinasa (ASK), concretamente CyrN. Las ASK son enzimas que catalizan la formación de PAPS, que es el donante de sulfato para las sulfotransferasas. Se propone que CyrJ sulfata la cilindrospermopsina en C12 mientras CyrN crea el conjunto de PAPS requerido para esta reacción. El

cribado de cepas productoras y no productoras de cilindropermopsina reveló que los genes de sulfotransferasa sólo estaban presentes en las cepas productoras de cilindropermopsina, afirmando adicionalmente la implicación de esta agrupación completa en la biosíntesis de cilindropermopsina (Figura 7). El gen *cyrJ* podría ser por lo tanto un buen candidato para una sonda de toxina, ya que es más único que los genes NRPS y PKS y presumiblemente tendría menos reactividad cruzada con otras agrupaciones de genes que contienen estos genes, que son comunes en cianobacterias. La reacción de adaptación final es realizada por *Cyrl*. Una búsqueda Fugue y un Psi-Blast iterado revelaron que *Cyrl* es similar a una hidroxilasa que pertenece a la superfamilia de oxigenasas dependientes de 2-oxoglutarato y Fe(II), que incluye la subunidad alfa de Prolil 4-hidroxilasa de mamíferos que cataliza la hidroxilación del colágeno. Se propone que *Cyrl* cataliza la hidroxilación de C7, un residuo que, junto con el anillo de uracilo, parece conferir mucha de la toxicidad de cilindropermopsina. La hidroxilación en C7 por *Cyrl* es probablemente la etapa final en la biosíntesis de cilindropermopsina.

Ejemplo 13: Transporte de la toxina CYR

Parece que la cilindropermopsina y otras toxinas cianobacterianas se exportan fuera de las células productoras. La agrupación de genes de cilindropermopsina contiene un ORF denominado *CyrK*, cuyo producto es lo más similar a proteínas de extrusión multi-fármaco y de compuestos tóxicos (MATE) dirigidas por el ión sodio de la familia NorM. Se postula que *CyrK* es un transportador para cilindropermopsina, tomando como base esta homología y su localización central en la agrupación. La expresión heteróloga y caracterización de la proteína se están realizando actualmente para verificar su posible papel en la exportación de cilindropermopsina.

Ejemplo 14: Regulación transcripcional de la agrupación de genes de la toxina

Se ha mostrado que la producción de cilindropermopsina es mayor cuando se elimina el nitrógeno fijado del medio de crecimiento (Saker et al. (1999) J. Phycol 35:599-606). Flanqueando la agrupación de genes de cilindropermopsina están homólogos del gen "hyp" implicados en la maduración de hidrogenasas. En la cianobacteria *Nostoc PCC73102* están bajo la regulación del regulador de nitrógeno global *NtcA*, que activa la transcripción de genes de asimilación del nitrógeno. Es plausible que la agrupación de genes de cilindropermopsina esté bajo la misma regulación, ya que está localizada completamente en la agrupación de genes "hyp" en *C. raciborskii* AWT205, y no pudo identificarse ninguna región promotora obvia en la agrupación de genes de cilindropermopsina.

Finalmente, la agrupación de genes de cilindropermopsina también incluye un ORF en su extremo 3' designado *CyrO*. Por homología, codifica una proteína hipotética que parece poseer un casete de unión a ATP, y es similar a proteínas con repetición WD, que tienen diversos papeles reguladores y de transducción de la señal. *CyrO* también puede tener un papel en la regulación transcripcional y unión a ADN. También muestra homología con las proteínas de la familia AAA que realizan frecuentemente funciones semejantes a chaperona y asisten en el ensamblaje, operación, o desensamblaje de complejos proteicos. La comprensión más profunda del papel de *CyrO* está dificultada debido a la baja homología de secuencia con otras proteínas en las bases de datos.

LISTADO DE SECUENCIAS

<110> NewSouth Innovations Pty Ltd.

5 <120> Detección de Organismos Cianotóxicos

<130> 852090

<160> 186

10 <170> PatentIn version 3.2

<210> 1

<211> 37606

15 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 1

atgatcccag	ctaaaaaagt	ttatttttta	ttgagtttag	caatagttat	ttcacccttt	60
ttatccatga	ttgtgggtat	ttacgaaaat	attaaattta	gggtattatt	tgatttggtg	120
gtcagggcac	taatgggtgt	tgactgcttc	aatatcaaaa	aacatcgggt	caaaattagt	180
cgtcaattac	ctctacgttt	atctattgga	cgtgagaatt	tagtaattat	gaaggtagag	240
tctgggaatg	tcaatagtgc	tattcaaatt	cgtgattact	atcccacaga	atttcccgtg	300
tccacatcta	acctgatagt	taaccttccc	cctaatacata	ctcaggaagt	aaagtacacc	360
attcgacctg	atcaacgggg	agaatttttg	tggggaaata	ttcaagttcg	acagctggga	420
aattgggtctc	taggggtggg	caattggcaa	attccccaaa	aaactgtggc	taagggtgat	480
cctgatttgt	taggactcag	atccctcgct	attcgtttta	ccctacaatc	ttctggatct	540
atcactaaat	tgctgcaacg	gggaatggga	acggaatttg	ccgaactccg	taattactgc	600
atgggggatg	atctacgggt	aattgattgg	aaagctacag	ctagacgtgc	ttatggaaat	660
ctgagtcctc	tagtaagagt	tttagagcct	caacaggaac	aaactctgct	tatattatta	720
gatcgtggta	gactaatgac	agctaattga	caagggttaa	aacgatatga	ttggggttta	780
aataccacct	tgtctttggc	attagcagga	ttacataggg	gcgatcgcgt	aggagtaggg	840
gtatttgact	ccagctgca	tacctggata	cctccagagc	gaggacaaaa	tcacttcaat	900
cggcttatag	acagacttac	acctattgaa	ccagtgttag	tggagtctga	ttattttaat	960
gccattacct	atgtagtaaa	acaacagact	cgtagatctc	tagtagtggt	aattactgat	1020
ttagtctgct	ttactgcttc	ccatgaaact	ctagttagcg	tgtgtaaatt	agtgcctcga	1080
tatctacctt	tttgtgtaac	actcagggat	cctgggattg	ataaaatagc	tcataatttt	1140
agtcaagact	taacacaggc	ttataatcga	gcagtttctt	tggacttgat	atcacaaaga	1200
gaaattgctt	ttgctcagtt	gaaacaacag	ggagtttttg	tgttgatgac	accagcaaat	1260
caaatttccg	agcagttggt	agaaaggtac	ttacaaaatca	aagccaaaaa	tcagatttga	1320
ctccctgctg	agataattga	gaacttctgg	aaagaaatagc	ccaataaact	cgacaaagaa	1380
cgtgggttag	agtcttttaa	agagtctatc	atgccgaatc	atatttttaac	agaagagcga	1440
tcgctcttcc	taagggatag	agtctgaaag	ccacttcaac	ggacgataat	gcaactcttg	1500
ttccagctgg	agtgcggaga	attaccacat	ccgaaataga	caaaaagaaa	taattggagt	1560
taagaagata	agtacataaa	tagtgataat	atacaaaact	agtcagcacg	gattaaattt	1620
actaatgata	gatacaatat	cagtactatt	aagagagtgg	actgtaattt	cccttacagg	1680
tttagccttc	tggctttggg	aaattcgtct	tcccttccat	caaattgaaat	acaaagctaa	1740
attcttcaag	gaattgggat	gggcgggaat	atcattcgct	tttagaaatg	tttatgcata	1800
tgtttctgtg	gcaattataa	aactattgag	ttctctattt	atgggagagt	cagcaaat	1860
tgcaggagta	atgtatgtgc	ccctctggct	gaggatcatc	actgcatata	tattacagga	1920
cttaactgac	tatctattac	acaggacaat	gcatagtaat	cagtttcttt	gggtgacgca	1980
caaatggcat	cattcaacaa	agcaatcatg	gtggctgagt	ggaaacaaag	atagctttac	2040
cggcggactt	ttatatactg	ttacagcttt	gtgggttcca	ctgctggaca	ttccctcaga	2100
gggtatgtct	gtagtggcag	tacatcaagt	gattcataac	aattggatac	acctcaatgt	2160
aaagtggaa	tcctgggttag	gaataattga	atggatttat	gttacgcccc	gtattcacac	2220
tttgcatcat	cttgatacag	ggggaagaaa	tttagtctct	atgtttactt	tcacgaccg	2280
attattttgga	acctatgtgt	ttccagaaaa	ctttgatata	gaaaaatcta	aaaatagatt	2340
ggatgatcaa	tcagtaacgg	tgaagacaat	tttgggtttt	taatagactt	gggttctaa	2400
tggaatggac	ggaaaaaatg	gcgggttacc	gcattcttaa	tatatcctct	ttttggggtt	2460
gagattttgga	taaagcggct	tgtactctgt	cattattcaa	atagccatgg	cgttgcatat	2520
ttgcgggatg	atttaagatt	ttctccta	ttgaaaaatt	tctctttag	gacgattgag	2580
aagcactcgc	gagattgcat	tattaataaa	acctgatag	tcaccccaa	cttattgcag	2640
aaaaactttt	ttctcttagg	taataaatta	gtagtttaat	tgaaaagcat	agcatctctt	2700

20

ttgacttgga	ataacaaaat	gtctttacgat	gtagtctagc	taaatagtga	cgcaaacgac	2760
tgttttctcc	ctcaactcta	gtcattgatg	ttttactaat	aatttgggtct	ccatcgggaa	2820
taaattttgg	gtaaacttta	tagccatccg	taatccaaaa	ataggatttc	caatgctcta	2880
tctttttcca	taatttggca	aatgttttgg	cactttctatc	tcccactaca	tattgaataa	2940
ttcccgaacg	tttgttatct	acaactgtcc	agaccatat	cttgtttttt	tttaccaata	3000
aatgtttcca	actcatccag	ttgacaaaact	tcaggtgttt	gggaattatt	attactatct	3060
gataactgac	gacctagctt	tttgacccaa	cgaatgactg	tattgtgatt	tacttttagtc	3120
attctttcaa	ttgccttaaa	tccattccca	tttacatata	tggttaaaca	tgcttccttt	3180
acttcttggg	aataacctct	aggagaataa	gattcaataa	attgacgacc	acaattcttg	3240
cattgataat	tttgttttcc	ccttctctgg	ccattttttc	taatattatt	ggaatcacag	3300
tttgaacagt	tcactttgat	ttcttctctg	cggcgatcgc	ctgctaataa	ttcttccct	3360
tattatacat	catcccgtgc	agggtgcaacg	cccaaatagc	catagtttat	gatcgggtatc	3420
gaattcgcta	ttgttttttc	tgccatatcc	cttacctaag	atgggacgat	attcgctcat	3480
aataccactg	tcaattagat	catcagcaac	atggtgagtg	tatcctgacg	accatcgata	3540
tggccaccaa	gatcactagc	tacccactg	ggcaacaatt	cgagtaaaag	cgagtagccc	3600
tactgtagca	ttgaaaccaa	ccaagtttga	agttaaatac	ctaaaattat	gacctcattt	3660
tcattttctag	acgttcagca	acgggcatta	actcacgtat	cagatcaaag	tttctctacgt	3720
tccgtctcat	ccagtctaat	aagaattttt	ctccttcatc	tagcttacct	ttatcatcaa	3780
caaaaaccaa	ctgctcgcac	caatctacaa	atccggaatt	agtcactctca	tagactaaaa	3840
tgatgggagg	aaagtgtgcg	aatcccat	tttcaatgac	ttccatacaa	accagcttaa	3900
atacttgttc	gtttgtcaat	tcattagaca	taaagaattt	tcctttaatc	aattctgttt	3960
ctaactctac	cacagagtaa	taactcttgg	tctggaacat	aaattattct	gtttttatca	4020
atgcgtaagt	cataacttat	tacttgacgg	agttgcaggg	gcataacctta	acttgacctt	4080
gggagcgata	gaagaaagga	aggcttcagt	gacgggtctt	tgactaatcc	cagtttccac	4140
ttcaactaaa	acagcatcac	aaatgtcgaa	tagtgattga	gaatatctat	tcataattcat	4200
gaaagtccaga	cgagattcca	tcggagacat	ggatgaatta	aaggcagcgt	tttcagcgta	4260
tcgacctgta	aatatatttc	cgtgggaatc	ttttaacgct	accctgcaa	aatttttcgt	4320
gtagggagca	taactttgat	tggcagcgga	tagagcagca	agcacaacat	catcggtaga	4380
ataggtctcc	agatcatgaa	atactgtttg	cattaatcca	cctgtgagtc	ctagatccgc	4440
tgggtccaaa	ggctcgggta	gaaaatgtgg	gagtttattt	gaggtataag	tttgctcagg	4500
ctgtgattca	ttagacttca	caagaagaac	aaaattttga	tttacagttg	ccatctcgta	4560
taaaaattgt	cggcagtatc	cacatggtgc	ttcgtggatt	gctaattgctt	gtaaaccggt	4620
ttctccgtgc	aaccacgcat	ttatgggtgc	ggattgttct	gcgtgaactg	agaaactaag	4680
tgccctgcct	acaaattcca	tgctcggcacc	aaaaataaga	gttccagaac	ccagttgatt	4740
cttagattgt	ggtttaccac	gagcgatcgc	ccctacataa	aactgcgata	ttggtaccct	4800
agcataagtt	gcggctacgg	gtagtaattg	aatcattaac	gtactaatat	tagtaccag	4860
tcgatcaatc	caagatgcga	caacacttga	gtcaattaca	gcatgttggg	caagaattgt	4920
ccttaactct	gattgaatgg	aacgtggaac	cttggcaatc	gcctgttcta	atgctacatg	4980
gggtcatttg	gttattcttg	gacagagaga	taaagatata	ttagttttta	tgaatcaatt	5040
tcccacttaa	tgcttgagta	tgttttctct	ctgcttacia	ggcaaagctt	tccttttttg	5100
tagcaaatcc	caactgctt	tgagagattt	aattgcttgg	tctatctcct	cttcggtatt	5160
ggcggctgta	atcgaaaacc	ttaaagcact	tttattttaa	ggtacgattg	gaaaaatagc	5220
aggagtaatt	aaaataccat	attcccaaag	gagttgacac	acatcaatca	tgtgttgagc	5280
atctcccact	aacacgccta	cgatgggaac	gtaaccatag	ttatccactt	cgaatccaat	5340
ggctcttgct	tgtgtaacca	atttgtgagt	taggtgataa	atttgttttc	ttactgctc	5400
cccctcctga	cgattcacct	gtaatccggc	taaggcaactt	gccaaactcg	caacaggaga	5460
aggaccagaa	aatatggcag	tccaagcgtt	gcggaagttg	gttttgatcc	ggcgatcgcc	5520
acaagttaag	aatgctgcgt	aagaagaata	ggctttggac	aaaccagcta	catagatgat	5580
attatcctct	gcaaacgcga	ggtcaaaaata	attcaccatc	ccgtttcctt	tgtaaccgta	5640
aggcatatcg	ctgctgggat	tttcgcccac	aatgccaaaa	ccatgagcat	catccatgta	5700
aattaaaggca	ttgtactctt	ttgccagatg	cacgtaagct	ggcagatcgg	gaaaatctgc	5760
cgacatggaa	tacacgccat	caatgacaat	aatctttact	tgttcaggcg	gatattttgc	5820
tagtttttgc	gctaaatcgt	tcaaatcatt	atgtcgatat	tggatgaact	gggctccttt	5880
gtgctgagcc	agacagcacg	cttcataaat	acaacgatgt	gcagctatgt	caccaaagat	5940
gacaccatta	ttcccagtta	atagtgttaa	aattcctatc	tgaagcagtg	ttacagctgg	6000
aaatactaaa	acatcaggta	cgcctaaaag	tttggacaat	tcttctcca	attcctcata	6060
aattgctggg	gaagcaacaa	gccgagtcca	gcttggatgt	gtgccccatt	tatccaaagc	6120

tggtggaatt	gcttccttaa	cttttggatg	caagtcaaga	cctaaatagt	tgcaagaagc	6180
aaagtctatc	acccaatgtc	cgatcaattag	caccttgcca	ccttggtgtt	ctgtgacgac	6240
tcttgtgact	tgaggaaattt	tttggttggt	aactacgttt	tccagagtgt	tgatttcgtt	6300
ggctgagtca	acaggtggag	ctagatcaga	ttgtttctct	tgtaccactt	ggttttggaa	6360
ataagtgatg	atggcagttg	gagtgttctt	ttgtaaaaag	aacgttccag	acagattgat	6420
ccctaaacgt	tcctctagga	gcgtttgcag	ttctaataaa	tctaaagaat	ctaattcccat	6480
atccagcagt	ttttgttgtg	gagcgtagcg	tgcctgacgt	tggaaccca	ttacttttaa	6540
gatgcattct	ttaacgagat	ccgctacagt	tttgttttcc	ttagttgcag	atgttgcttt	6600
tggtaccaat	gaaccaattg	ctgagttaat	atacggtcct	ttgcgatcac	cagggcagtg	6660
caaagcactg	tcgcgcaggt	tatattcaat	caaaataccc	atgccgagat	tatctgtatc	6720
ttccggacga	taattagcaa	taattcccct	aatttcggct	cctcccgcac	catggaaacc	6780
cacaattgga	tccagaagct	gtcgttgctc	attgtgtagc	tttaataact	ccatcatcgg	6840
catttgggaa	taattgacat	aatttcgaca	gcgagttaca	cccaccacgc	tctcaatgcc	6900
gcctttcagg	gtacagtagt	aaagcataaa	gtcccgaat	tcatttccta	acccccgcgc	6960
ctgaaactca	ggtagaatat	ttagtgcgag	cagttgaata	actgaccctt	ggggagtatg	7020
taacgtcggc	acttgcgcac	atttttacatt	ctctaagccc	tcagtgtctg	taattgtttg	7080
ggaataaata	gcaccaataa	tttgatcttc	tataatcagc	actaaattac	cttgcggttt	7140
tagctcaagt	cttcgccgaa	tttcatgagt	agatgcccg	aaattttctg	gccaacactt	7200
gacctccaag	tcaactaagg	caggtaaatc	tgacaaatag	gcagactaa	ttttgtgaagg	7260
tcttttctcg	aagtaattaa	gcgtaatgcg	agtaaaagga	aatgtttttg	ggtatctttt	7320
agaaagctct	agttttggaa	atagacctac	ttgtgcagca	gacatgagaa	aaacctcagc	7380
ttccacaaga	tactgctgag	aaaatccctg	aaacgcacgc	aaatgtaagt	tttcgctttt	7440
gtctaaaaac	tgatagacta	cccttggttc	caaacaatgg	acctccaaaa	tcattaaacc	7500
gtgtttattg	accacttgag	accatctttc	taagtgttcc	accaaacttt	gcaccataac	7560
atgaggagga	ataagctctc	cttgatcatc	gacacagact	gattggtaag	gtaagtgagc	7620
acgttcttcc	aattcgtttc	tttcttgagg	aggaataaag	agacgatcat	ggtcgaggaa	7680
cgaacggatg	tgaggatata	tttcgggagc	atgaatgcc	tgagcttcta	aagaacgcac	7740
catttgttct	gggttcccaa	tatctccctg	taaaactaag	tggggaaggc	tagcaagggt	7800
gcgtgtggta	gcttttaaa	aagcttctgt	ataatctaca	cctataagac	gcaggggata	7860
ctgttcgagt	gcttttcccc	tagcagactt	aaattgaatg	gtttcccaga	ctcgtttcag	7920
gagagttcca	tcgcccacac	ccatgtcagt	aatgtatttg	ggtgtttctt	ctaattggcaa	7980
ctgattgaat	actgagagga	tactttcttc	taaatcggca	aaatatttct	ggtgttgaaa	8040
tccactcccg	atcacgttaa	gggtgcgac	aatgtgcctt	tcgtgaccgg	aagcatctct	8100
ttggaatacg	gagagacaat	tgccaaaaca	tacatcatga	atgcgggaca	acataggagt	8160
gtaggacgcc	actatggctg	tattcaaggc	tcgtctctcc	ataaatcgac	caagttcggg	8220
tatggtcaaa	cgacctgctg	taaggtcagc	ccagccaagg	tgagagaaata	acttacccaa	8280
ctcttcttgc	actgttgagc	ttaatgagga	gagcaaaagg	ttgtcctccg	aatctgcaag	8340
caagttgtgt	ttgtgcagtg	ccagcaggag	tgggatgacc	agtaatccat	ctaaaaaatc	8400
tgccattagg	ggattgtcca	ggttccacaa	ttggcaagaa	cgctcaatcc	atcttcccag	8460
caaatttcct	tgtttccctt	ctaaataaga	ctgaattggg	aggttgtaca	attgaagaat	8520
gtcttccgaa	attttgttgt	gaatcgctgc	ttctgcgggt	agagagtatt	taagctcctt	8580
atttcgggaa	agccaatgta	aagactcgag	catcctcaaa	gcaacttgaa	aatgtccgct	8640
gtagctccc	agatgttcca	ccatttggtt	taaagagaga	ggactttcat	cggcgagtaa	8700
ttcaaaaaca	cctttttctc	gacacgcaag	aataacggga	accgccacaa	agccgtgagt	8760
ataacgatta	atcttttgta	acatttagac	gattattgat	taatttatga	ggaatgcatt	8820
tttagtgcat	accacgagat	tttgattgtc	tcagaagttg	tgtgaaaaag	caagacaagt	8880
agaccaaaaa	aataagctaa	ataagtgtag	tagcaataaa	aagacgaatc	gcaattgtac	8940
gtgtcttgac	taacaagcca	agtctctcta	gataataatc	gccctctacc	agttgcgtaa	9000
gtcccattgt	tgttttaaac	tttaattgct	aattaaacag	ttatcaaatc	ctgttcataa	9060
cggatattta	cagcaatttt	cggttatata	aaattgcata	tactgtaagt	aatagcagaa	9120
aattaattta	ggtaggaaaa	tgttgaaaga	ttttcaaccg	tttttaatca	gaacactagc	9180
attcgtattc	gcatttggtg	ttttcttaac	cactggagtt	ggcattgcta	aagctgacta	9240
cctagttaaa	ggtggaaaga	ttaccaatgt	tcaaaatact	tcttctaacg	gtgataatta	9300
tgccgttagt	atcagcgggt	ggtttgttcc	ttgcgcagat	agagtgatta	tcctaccaac	9360
ttcaggagtg	ataaatcgag	acattcatat	gcgtggctat	gaagccgcat	taactgcact	9420
atccaatggc	tttttagtag	atatttacga	ctactctggc	tcttcttgca	gcaatggtgg	9480
ccaactaact	attaccaacc	aattaggtaa	gctaactcagc	aattagggtg	tatcatgata	9540

agatgaagta	gtttaacat	ggcaccacca	gccaaaaact	ttttaacgct	aggggtgaac	9600
agttatgggt	gtggaatgta	ggttgatcc	agtgcataa	acagccataa	ttttagtata	9660
agcaaacact	aagattggag	aattcatgga	aacaacctca	aaaaaattta	agtcagatct	9720
gatattagaa	gcacgagcaa	gcctaaagtt	gggaatcccc	ttagtcattt	cacaaatgtg	9780
cgaaacgggt	atttatacag	cgaatgcagt	catgatgggt	ttacttggta	cgcaagtttt	9840
ggccgcccgt	gctttgggcg	cgctcgcttt	tttgacctta	ttatttgcct	gccatggtat	9900
tctctcagta	ggaggatcac	tagcagccga	agcttttggg	gcaaataaaa	tagatgaagt	9960
tagtcgtatt	gcttccgggc	aaatatggct	agcagttacc	ttgtctttac	ctgcaatgct	10020
tctgctttgg	catggcgata	ctatcttgct	gctattcggt	caagaggaaa	gcaatgtgtt	10080
attgacaaaa	acgtatttac	actcaatttt	atggggcttt	cccgtgcgc	ttagtatttt	10140
gacattaaga	ggcattgcct	ctgctctcaa	cgttccccga	ttgataacta	ttactatgct	10200
cactcagctg	atattgaata	ccgccgccga	ttatgtgtta	atattcggta	aatttggctt	10260
tcctcaactt	ggtttggctg	gaataggctg	ggcaactgct	ctgggttttt	gggttagttt	10320
tacattgggg	cttatcttgc	tgattttctc	cctgaaagtt	agagattata	aacttttccg	10380
ctacttgcat	cagtttgata	aacagatctt	tgtcaaaatt	tttcaactg	gatggcccat	10440
gggggtttcaa	tgggggggcg	aaacggcact	atttaacgtc	accgcttggg	tagcagggtat	10500
tttaggaacg	gtaacattag	cagcccatga	tattggcttc	caaacggcag	aactggcgat	10560
ggttatacca	ctcggagtcg	gcaatgtcgc	tatgacaaga	gtaggtcaga	gtataggaga	10620
aaaaaacctt	ttgggtgcaa	gaagggtagc	atcgattgga	attacaatag	ttggcattta	10680
tgccagtatt	gtagcacttg	ttttctgggt	gtttccatat	caaattgccg	gaattttatt	10740
aaatataaac	aatcccgaga	atatcgaaag	aattaaagaa	gcaactactt	ttatcccctt	10800
ggcgggacta	ttccaaatgt	tttacagtat	tcaaataatt	attgttgggg	ctttggctcg	10860
tctgcgggat	acatttggtc	cagtatcaat	gaacttaatt	gtctgggggc	ttggattggc	10920
aggaagctat	ttcatggcaa	tcattttagg	atgggggggg	atcgggattt	ggttggctat	10980
ggttttgagt	ccaactcctc	cggcagttat	tttaactgtt	cgtttttata	gagtgtattg	11040
caatcttctt	ccaacagtg	atgatagtgt	acagaatgcg	tctgttacta	ctctaggctg	11100
agaaaagcta	tatgaccaat	caaaaataacc	aagaattaga	gaacgattta	ccaatcgcca	11160
agcagccttg	tccggtcaat	tcttataatg	agtgggacac	acttgaggag	gtcattgttg	11220
gtagtgttga	aggtgcaatg	ttaccggccc	tagaaccaat	caacaaatgg	acattcccct	11280
ttgaagaatt	ggaatctgcc	caaaagatac	tctctgagag	gggaggagtt	ccttatccac	11340
cagagatgat	tacattagca	cacaaagaac	taaatgaatt	tattcacatt	cttgaagcag	11400
aaggggtcaa	agttcgtcga	gttaaaccctg	tagattttct	tgtccccttc	tccacaccag	11460
cttggaaggt	aggaagtggg	ttttgtgccc	ccaatcctcg	cgatgttttt	ttgggtgattg	11520
ggaatgagat	tattgaagca	ccaatggcag	atcgcaaccg	ctattttgaa	acttgggcgt	11580
atcgagagat	gctcaaggaa	tatttttcagg	caggagctaa	gtggactgca	gcgccgaagc	11640
cacaattatt	cgacgcacag	tatgacttca	atttccagtt	tcctcaactg	ggggagccgc	11700
cgcgtttctg	cgttacagag	tttgaaccga	cttttgatgc	ggcagatttt	gtgcgctgtg	11760
gacgagatat	ttttgggtcaa	aaaagtcatg	tgactaatgg	tttgggcata	gaatgggttac	11820
aacgtcactt	ggaagacgaa	taccgtattc	atattattga	atcgcatgtt	ccggaagcac	11880
tgcacatcga	taccacctta	atgcctcttg	cacctggcaa	aatactagta	aatccagaat	11940
ttgtagatgt	taataaattg	ccaaaaatcc	tgaaaagctg	ggacattttg	gttgcacctt	12000
accccaacca	tatacctcaa	aaccagctga	gactggtcag	tgaatgggca	ggtttgaatg	12060
tactgatgtt	agatgaagag	cgagtcattg	tagaaaaaaa	ccaggagcag	atgattaaag	12120
cactgaaaga	ttggggattt	aagcctattg	tttgccattt	tgaaagctac	tatccatttt	12180
taggatcatt	tcactgtgca	acattagacg	ttcgccgacg	cggaactctt	cagtcctatt	12240
tttaagattt	atttcogatta	tcctttatcc	tgatcatcca	gagtgtataag	agcattacaa	12300
ctaggagaca	attatgacaa	ctgctgacct	aatcttaatt	aacaactggg	acgtagtctg	12360
aaaggtggaa	gattgtaaac	caggaagtat	caccacggct	cttttatttg	gagttaagtt	12420
ggtactatgg	cgagtcgtg	aacagaattc	ccccatacag	atatggcaag	actactgcc	12480
tcaccgaggt	gtggctctgt	ctatgggaga	aattgttaat	aatacttttg	tttgtccgta	12540
tcacggatgg	agatataatc	aagcaggtaa	atgcgtacat	atcccggctc	accctgacat	12600
gacaccccca	gcaagtgcct	aagccaagat	ctatcattgc	caggagcgat	acggattagt	12660
atgggtgtgc	ttagggtgatc	ctgtcaatga	tataccttca	ttacccgaat	gggacgatcc	12720
gaattatcat	aatacttgta	ctaaatctta	ttttattcaa	gctagtgcgt	ttcgtgtaat	12780
ggataatttc	atagatgtat	ctcattttcc	ttttgtccac	gacgggtgggt	taggtgatcg	12840
caaccacgca	caaattgaag	aatttgaggt	aaaagtagac	aaagatggca	ttagcatag	12900
taaccttaaa	ctccagatgc	caaggtttaa	cagcagtaac	gaagatgact	catggactct	12960

ttaccaaaagg	attagtcac	ccttggtgca	atactatatt	actgaatcct	ctgaaattcg	13020
gactgcggat	ttgatgctgg	taacaccgat	tgatgaagac	aacagcttag	tgcaaatgtt	13080
agtaacgtgg	aaccgctccg	aaatattaga	gtcaacggta	ctagaggaat	ttgacgaaac	13140
aatagaacaa	gatattccga	ttatacactc	tcaacagcca	gcgcgtttac	caactgttacc	13200
ttcaaagcag	ataaacatgc	aatgggtgtc	acaggaata	catgtaccgt	cagatcgatg	13260
cacagttgcc	tatcgctgat	ggctaaagga	actgggcgtt	acctatggtg	tttggttaatt	13320
tcagggttgt	tggtatctgg	ataggtatgg	ttttgagtcc	actgctatct	ggagggattt	13380
taatggttgg	tttttatcaa	cagcttgcca	ataagtatta	ctaatagtga	tgatggggaa	13440
gagaatcaaa	ctatactcac	caacaagggtg	ttaaaatgca	gatcttagga	atttcagctt	13500
actaccacga	tagtgctgcc	gcgatggtta	tcgatggcga	aattggttgc	gcagctcagg	13560
aagaacgttt	ctcaagacga	aagcacgatg	ctgggtttcc	gactggagcg	attacttact	13620
gtctaaaaca	agtaggaacc	aagttacaat	atatcgatca	aattgttttt	tacgacaagc	13680
cattagtcaa	atttgagcgg	ttgctagaaa	catatttagc	atatgcccc	aagggatttg	13740
gctcgtttat	tactgctatg	cccgtttggc	tcaaagaaaa	gctttaccta	aaaacacttt	13800
taaaaaaaga	attggcgctt	ttggggaggt	gcaaagcttc	tcaattgcct	cctctactgt	13860
ttacctcaca	tcaccaagcc	catgcggccg	ctgctttttt	tcccagtcct	tttcagcgtg	13920
ctgccgttct	gtgcttagat	ggtgtaggag	agtgggcaac	tacttctgtc	tggttgggag	13980
aaggaaataa	actcacacca	caatgggaaa	ttgattttcc	ccattcccct	ggtttgcttt	14040
actcagcgtt	tacctactac	actgggttca	aagttacttc	aggtgagtag	aaactcatgg	14100
gttttagcacc	ctacggggaa	cccaaatatg	tggaccaaat	tctcaagcat	ttgttggatc	14160
tcaaagaaga	tggtactttt	aggttgaata	tggactactt	caactacacg	gtggggctaa	14220
ccatgaccaa	tcataagttc	catagtatgt	ttggaggacc	accacgccag	gcggaaggaa	14280
aaatctccca	aagagacatg	gatctggcaa	gttcgatcca	aaagtgact	gaagaagtca	14340
tactgcgtct	ggctagaact	atcaaaaaag	aactgggtgt	agagtatcta	tgtttagcag	14400
gtgggtgcgg	tctcaattgc	gtggctaacg	gacgaattct	ccgagaaagt	gatttcaaag	14460
atatttggat	tcaacccgca	gcaggagatg	ccggtagtgc	agtgggagca	gcttttagcga	14520
tttgcatga	ataccataag	aaacctcgca	cttcaacagc	aggcgatcgc	atgaaagggt	14580
cttatctggg	acctagcttt	agcgaggcgg	agattctcca	gtttcttaat	tctgttaaca	14640
taccctacca	tcgatgcgtt	gataacgaac	ttatggctcg	tcttgacaga	atttttagacc	14700
agggaaatgt	tgtaggctgg	ttttctggac	gaatggagtt	tggtccgcgt	gctttgggtg	14760
gccgttcgat	tattggcgat	tcacgcagtc	caaaaatgca	atcggtcatg	aacctgaaaa	14820
ttaaatatcg	tgagtccttc	cgtccatttg	ctccttcagt	cctggctgaa	cgagtcctcg	14880
actacttcga	tcttgatcgt	cctagtcctt	atatgtcttt	ggtagcacia	gtcaaagaga	14940
atctgcacat	tccatgaca	caagagcaac	acgagctatt	tgggactcgag	aagctgaatg	15000
ttcctcgctc	ccaaattccc	gcagtcactc	acgttgatta	ctcagctcgt	attcagacag	15060
ttcacaaaga	aacgaatcct	cgttactacg	agttaattcg	tcatttttag	gcacgaactg	15120
gttgtgctgt	cttgggtcaat	acttcgttta	atgtccgcgg	cgaaccaatt	gtttgtactc	15180
ccgaagacgc	ttatcgatgc	tttatgagaa	ctgaaatgga	ctatttggtt	atggagaatt	15240
tcttggttgg	caaatctgaa	cagccacggg	gaaatagtga	tgagtcatgg	caaaaagaat	15300
tcgagttaga	ttactttatg	agtgaatttt	tcccacaaaa	aagtggtaaa	ttaaagatgg	15360
aacagataaa	agaacttgac	aaaaaaggat	tgctgtgagt	tggactgatt	ggcgggttcta	15420
tagtgggcgt	tttatctggc	tttttactgc	cagttatacg	ccatcattcc	ttatcagtta	15480
tcccttgggt	tggtgctgga	tttctctgga	tttgggcaat	aatcgacact	acgacttta	15540
gttttattta	ccaaatatgg	atgaggattg	gacttggttt	aggatggata	caaacacgaa	15600
ttattttggg	agttttatlt	tatataatga	tcacaccaat	aggattcata	agacggctgt	15660
tgaatcaaga	tccaatgacg	cgaatcttcg	agccagagtt	gccaaacttat	cgccaattga	15720
gtaagtcaag	aactacacaa	agtatggaga	aaccattcta	atgctaaaag	acacttgga	15780
ttttattaaa	gacattgccg	gatttattaa	agaacaaaaa	aactatttgt	tgattcccct	15840
aattatcacc	ctgggtatcct	tgggggcgct	gattgtcttt	gctcaatcct	ctgcgatcgc	15900
acctttcatt	tacactcttt	tttaaattgc	catattatga	gtaacttcaa	gggttcggt	15960
aagatagcat	tgatgggaat	attgattttt	tgtgggctaa	tctttggcgt	agcatttgtt	16020
gaaattgggt	tacgtattgc	cgggatcgaa	cacatagcat	tccatagcat	tgatgaacac	16080
agggggtggg	tagggcgacc	tcattgtttc	gggtgggtata	gaaccgaagg	tgaagctcac	16140
atccaaatga	atagtgtatg	ctttcgagat	cgagaacaca	tcaaggtaaa	accagaaaat	16200
accttcagga	tagcgctgtt	gggagattcc	ttttagagat	ccatgcaagt	accgttggag	16260
caaaatttgg	cagcagttat	agaaggagaa	atcagtagtt	gtatagcttt	agctggacga	16320
aaggcggaag	tgattaattt	tggagtgcac	ggttatggaa	cagaccaaga	actaattact	16380

ctacgggaga	aagtttggga	ctattcacct	gatatagtag	tgctagattt	ttatactggc	16440
aacgacattg	ttgataactc	ccgtgcgctg	agtcagaaat	tctatcctaa	tgaactaggt	16500
tcactaaagc	cgttttttat	acttagagat	ggtaatctgg	tggttgatgc	ttcgtttatc	16560
aatacggata	attatcgctc	aaagctgaca	tggtggggca	aaacttatat	gaaaaataaa	16620
gaccactcac	ggatttttaca	ggtttttaac	atggtacggg	atgctcttaa	caactctagt	16680
agagggtttt	cttctcaagc	tatagaggaa	ccgttattta	gtgatggaaa	acaggataca	16740
aaattgagcg	ggttttttga	tatctacaaa	ccacctactg	accctgaatg	gcaacaggca	16800
tggaagtcga	cagagaaact	gattagctca	atgcaacacg	agggtactgc	gaagaaagca	16860
gatttttttag	ttgttacttt	tgggcggtccc	tttcaacgag	aacctttagt	gcgtcaaaaa	16920
gaaatgcaag	aattgggtct	gactgattgg	ttttaccag	agaagcgaat	tacacgtttg	16980
ggtgaggatg	aggggttcag	tgtactcaat	ctcagcccaa	atttgcaggt	ttattctgag	17040
cagaacaatg	cttgccctata	tggttttgat	gatactcaag	gctgtgtagg	gcattggaat	17100
gcttttaggac	atcagggtagc	aggaaaaatg	attgcatcga	agatttgtca	acagcagatg	17160
agagaaagta	tattgcctca	taagcacgac	ccttcaagcc	aaagctcacc	tattacccaa	17220
tcagtgtatcc	aataaagaac	tgggcatcac	ttatgatgtt	tactaatttc	agttccggtg	17280
atgttaatgc	gtaactttta	ttactagtgt	taaagctgag	atatgacaaa	taccgaaaga	17340
ggatttagcag	aaataacatc	aacaggatat	aagtccagagc	ttagatcggg	ggcagcaggt	17400
agcctccaac	tggaatttcc	cttagtcctt	gtcgaaatat	gcggaacgag	tattaatgtg	17460
gtggatgtag	tcattgatggg	cttacttggg	actcaagttt	tggtctgtgg	tgcttgggtg	17520
gcgatcgctt	ttttatctgt	atcgaatact	tgttataata	tgcttttctg	gggggttagca	17580
aaggcatctg	aggcttttgg	ggcaaacaaa	atagatcagg	ttagtctgat	tgcttctggg	17640
caaatatggc	tggcactcac	cttgtctttg	cctgcaatgc	ttttgctttg	gtatatggat	17700
actatatattg	tgctatttgg	tcaagttgaa	agcaacacat	taattgcaaa	aacgtattta	17760
caactaatgg	tgtggggagg	tccggcgcca	gttgggtatt	tgatattaag	aggcattggc	17820
tctgctgtga	acgtccccc	attggtaact	gtgacgatgc	tagtagggct	ggtcttgaat	17880
gccccggcca	attatgtatt	aatgttcggg	aaatttgggc	ttcctgaact	tggttttagct	17940
ggaataggct	gggcaagtac	tttgggtttt	tggttaggtt	ttctagtggg	ggttgtcttg	18000
ctgattttct	ccccaaaagt	tagagattat	aaacttttcc	gctacttgca	tcagtttgat	18060
cgacagacgg	ttgtggaaat	ttttcaaact	ggatggccta	tggtttttct	actggggagt	18120
gaatcagtag	tattgagcct	caccgcttgg	ttaacaggct	atttgggaac	agtaacatta	18180
gcagctcatg	agatcgcgat	ccaaacagca	gaactggcga	tagtgatacc	actcggaatc	18240
gggaatggtg	ccgtcacgag	agtaggtcag	actataggag	aaaaaaaccc	tttgggtgct	18300
agaagggcag	cattgattgg	gattatgatt	ggtggcattt	atgccagtct	tgtggcagtc	18360
attttctggg	tgtttccata	tcagattgcg	ggactttatt	taaaaataaa	cgatccagag	18420
agtatggaag	cagttaagac	agcaactaat	tttctcttct	tgggcgggatt	attccaattt	18480
tttcatagcg	ttcaaataat	tgttgttggg	gttttaatag	ggttgcaagg	tacgtttatc	18540
ccattgttaa	tgaatttggg	aggctggggg	cttggcttgg	cagtaagcta	ttacatggga	18600
atcattttat	gttggggagg	tatgggtatc	tggttaggtc	tggttttgag	tccactcctg	18660
tccggactta	ttttaatggg	tcgtttttat	caagagattg	ccaataggat	tgccaatagt	18720
gatgatgggc	aagagagtat	atctattgac	aacgttgaag	aactctcctg	acgaacagat	18780
tgaattgcct	tggtcttgac	acttcgttaa	cctaagcatg	agagtatagg	ctatactctg	18840
ccgtgggtta	ctgagtgttg	tcctggatcg	aggacgcagc	ctggctgagc	aacaaaaaag	18900
actggaatct	tgacctgtca	atggttttta	ctgctagtgt	gaggctgggtg	tcagcagctt	18960
cgccatttct	gcgcctaaga	cttgacctag	ccataatatt	ttagtattat	gatgagcgat	19020
cttaatcaaa	ggcaaaaaat	ttacaattaa	tctattgtta	cattaatttt	gctcctcatt	19080
ctgttttaaat	tttcagtgac	attgtaatct	aaactcaaaat	gaaaaacaaac	aaacatatag	19140
ctatgtgggc	ttgtcctaga	agtcgttcta	ctgtaattac	ccgtgctttt	gagaacttag	19200
atgggtgtgt	tgtttatgat	gagcctctag	aggctccgaa	tgtcttgatg	acaacttaca	19260
cgatgagtaa	cagtcgtacg	ttagcagaag	aagacttaaa	gcaattaata	ctgcaaaata	19320
atgtagaaac	agacctcaag	aaagttagat	aacaattgac	tggaagttaa	ccggacggaa	19380
aattattctc	atttcaaaaa	atgataacag	gtgactatag	atctgaattt	ggaatagatt	19440
gggcaaaaaa	gctaactaac	ttctttttta	taaggcatcc	ccaagatatt	attttttctt	19500
tcgatataag	ggagagaaaag	acaggtatca	cagaaccatt	cacacaacaa	aatcttggca	19560
tgaaaacact	ttatgaagtt	ttccaacaaa	ttgaagttat	tacagggcaa	acacctttag	19620
ttattcactc	agatgatata	attaaaaacc	ctccttctgc	tttgaaatgg	ctgtgtaaaa	19680
acttagggct	tgcatattgat	gaaaagatgc	tgacatggaa	agcaaatcta	gaagactcca	19740
attttaaagta	tacaaaatta	tatgctaatt	ctgctctggg	cagttcagaa	ccttgggttg	19800

aaactttaag	atcgacacaaa	acattttctcg	cctatgaaaa	gaaggagaaa	aaattaccag	19860
ctcggttaat	acctctacta	gatgaatcta	ttccttacta	tgaaaaactc	ttacagcatt	19920
gtcatatttt	tgaatgggtca	gaacactgag	tttgatcgta	accggttcaga	ggggggatag	19980
aagcgcgatt	aggagatcc	aaaaataaaa	atatctagcc	gtctaacctc	tttattttca	20040
tcgattcttc	ttaccgttcc	ctattccctc	ccttcaccag	ttcgtttttg	ggtaggtgca	20100
agatctgagc	ctcccaccta	gggccgatct	ggcagtgcgc	gatcgccact	agcccatgga	20160
aaactagcac	tttttgggga	acagccaaaa	cctttattga	gtaagaattt	gaaaaagtgc	20220
aagttaagag	gcaatgacta	aaaatttttt	tctactcttt	tcaggataga	attccagttt	20280
ctagagccgt	tgtaacggta	catatcttga	tagtacgtat	cgatgaggta	ctcatttttcg	20340
tggagcatta	accagctttt	taactccgct	aatttctgct	ctcctttttc	tattaattct	20400
tgctcatcca	aatcatccct	gtccaactcc	tccctgtcca	actcccacat	agttttgttg	20460
gtatcttcga	caatcaagta	gtctccactt	tttagaccgt	tttcgtgaaa	atattcaact	20520
actcccaccg	cattagcatg	ggcatcttct	acgatcaacc	agggatgagc	aagcccagaa	20580
agcagttccg	acgacattat	tgcaccata	ttgttacaat	ccccctctaa	aaaatgaacg	20640
cgagagtcag	tttttgcctt	ctcgtcgagt	agggaaaagt	cgatatcgat	acagtagaca	20700
caaccttcta	tttgaacag	ttctaagtga	tcggctagcc	aaatcgcgct	gccaccgctt	20760
aatgctccta	tttcgattat	tgttttcggg	cgaagctcat	acaggagcat	tgaataaaga	20820
gctatttcgg	tgcacccttt	caggaagggt	atccctttcc	aagtgaacaa	atcgcgggtt	20880
gccaagagcg	ctctccaagc	tggcaactgga	atagcacatt	tatcttctct	ttcagaaatt	20940
ttggcaaac	gatttaggtt	gaaagggtgca	actttatagg	cggcttcttg	aacaaatttt	21000
tggaagctca	tctaattttc	ctcttaggtg	ttagaacatt	tgtaaaatct	tggcgatttt	21060
ttgttttctt	tcttgaatat	agcaaccgcc	aaggcggttt	gagcataaac	tggatgtagt	21120
ccccgtgttt	tacggttgag	acttaggtaa	agcggttttg	tttgtactct	cccattatct	21180
aaatagccgt	agtttatgat	cggtatccaa	ttcgctattg	ttttttctgc	catatcccca	21240
acctaagatg	cgacgatatt	cacccataat	gccactgtca	attaaatcat	cctcgttgac	21300
tgcaacattg	gtatgagatt	gcggcgcaac	atagagcgca	tccgcaggac	aatatgcttc	21360
acagatgaaa	caagtttgac	agtcttccctg	tcgggcgatc	gcaggcggtt	ggttgggaac	21420
tgcatacaag	acattggtag	ggcatacttg	gacgcaaa	ttacaattaa	tacagagttt	21480
atggctgaca	agctcgatca	tcatactgct	cctgctacaa	ctttaatact	ggggctgtgg	21540
tttaagtgg	taatactggt	ggtgtagcgc	tcgcatcctt	cacccaatcc	cgtctcacc	21600
aaagcctttc	taagccgccc	gtggcttgg	aataaagctg	atttggatcg	gtttcaggat	21660
agtctatg	aatatgttcg	ctacgcgttt	ccttgcgatg	taaagcgcta	aaatatgccc	21720
atcgtgctac	agacacaaga	gcagccgctc	gacgagaaaa	ttccagatcg	cgactgtat	21780
cttgtttcgg	gttcccttgt	acttgcgtgc	acagcatttc	taatttggcg	agggaatcca	21840
aaagtccctg	ctcacagcgc	aagtaattct	tctctaattg	gaacatctcg	gcttgtacac	21900
cgcggaaca	tgcctcgcta	tcgaatgttt	cggaaccagg	gtactgggaa	cgtaatccgg	21960
cttgacctgc	tggacgcaca	accggttcat	ggacatgagc	gccccaaactc	ttggcaaaag	22020
cggtgcacac	ttcccctgcc	cattgtcctg	tagagattgc	ccaagcagca	ttaggacct	22080
cacccccaga	agctatccca	gctaaaaact	cccgcgatgc	tgcattctcg	gcggcataca	22140
gtccaggaa	ttttgtacca	caactatcat	tcacaatccg	aattccacct	gtaccacgga	22200
ctgtaccttc	taaaaccagt	gttacaggta	ctcgttctgt	ataagggtca	atgccagctt	22260
ttttataggg	tagaaaggcg	atgaagtgg	acttttcaac	caatgcttgg	atttcaggtg	22320
tggctcgatc	caaacgagca	taaaccgggac	ctttcaggag	ggcattgggc	aggaacgatg	22380
gatcgcgacg	accattgata	tagccacca	gatcgttacc	tgcctcatcg	gtgtaactag	22440
cccagtaaaa	gggagcagcc	cttgtcactg	tggcattgaa	agcggtcgag	atggatatag	22500
gactggaagc	ttccatactg	gagagtctgc	cgccagcttc	caccgccatc	agcagtccat	22560
cgctgtatt	ggtattgcaa	cctaaagctt	tacttaggaa	tgcacaaccg	ccattcgcta	22620
gaactactgc	accagcgcga	acggtatagg	tgcgatgatt	ttgcctctgt	acacctctag	22680
ctccagccac	ggagccgtcc	tgggctaata	acagttctag	agccggactt	tggtcgaaaa	22740
tttgcacacc	cacacgcaac	aggttcttgc	gaagtaccgc	catatattcc	ggaccataat	22800
aactctggcg	cacggattcc	ccattttctt	tggggaaacg	atagcccaa	tcttccacta	22860
agggcaaaact	cagccaagct	ttttcaatta	cacgttcaat	ccaacgtaag	ttagcgaggt	22920
tatttctctt	gctgtaacat	tcggatacat	ctttctccca	attctctgga	gaagggtgcca	22980
tgacgctatt	gccactggca	gcagctgcac	cgctcgtacc	tagaaaacct	ttatcaacaa	23040
tgatgacttt	gacaccttgg	gctccagccg	cccatgctgc	ccatgcggcg	gcaggaccac	23100
caccaattac	cagcacgtca	gcagttaatt	gtagttcagt	gccgctatag	gctgtaagca	23160
attgcttttc	ctccttgttt	aaagtcaagt	tcatactttt	aattatcttc	tgcagtcggt	23220

cgaatcaaaa	tttcatttac	atttacatga	tcggggtgtg	tactgcata	aattatagct	23280
cttgcaatat	cctcactttg	taaaggtgtt	attgtactaa	gttgttcttt	actaagctgt	23340
ttcgtgatcg	ggtcagaaa	taagtcatta	aatggcgtat	cgactaaacc	tggctcaatg	23400
atggtaacgc	gaatgtgtgc	taaagatacc	tcctggcgta	atgcttctga	aagagcattg	23460
acgcctgatt	tggcagcact	ataaacgacc	gcaccggact	gcgctatcct	gccatcgaca	23520
gaagatatat	tgactatatg	accggatttt	tgggccttca	gaagaggcaa	aactgcgtgg	23580
atagcatata	aaactcccag	aacattcaca	tcgaatgctc	gcctccagtc	tgccgggattt	23640
ccagtatcaa	ttgcacaaaa	cacaccaatt	cctgcattat	tcacaaaaat	atctacatgt	23700
cctagctcaa	ccttgggtctt	ttggactaga	tgatttactt	gagattcgtc	tgtaatatct	23760
gtaacaatag	gcaatgcttg	accaccactg	gcttcaatcc	gttttgctag	tgcatgcaaa	23820
agctcagcac	gtcttgcggc	gatcgcaact	tttgccctct	ccgcagctaa	agcaaatgct	23880
gtagcctctc	caatcccaga	ggaagctcca	gtaataatcg	ccacttttcc	atccaattta	23940
cctgccatca	gtcactcctt	agttttcgtt	ttgctgggtc	aatatgtaat	aagtgcgttt	24000
tgtacttgat	tttgttcttt	ggtgattttt	atataggagc	gcataaagtg	cttagtgatc	24060
actttatttt	ttagtgcctt	tcaacttaaa	ttaacaaacc	ccataagtaa	cacctagtgt	24120
cttttagccat	cgacgatagg	caagtgtgca	tctatctgat	ggtacgtgga	tttcgtgtga	24180
aaacaattgt	gtattttatct	gctttggagt	taacagtggg	aaacgtaccg	gctgtgtgtgc	24240
atgtaagatc	cgaatatctt	gttctattgt	ttcgtcatat	tcagtttagca	tctttgactc	24300
taacgtttca	taccggttcc	acattatcaa	catagcgaat	acactatttt	cctcatcaat	24360
cgggtgtgatc	gtcattaaat	ccacaatcct	catttcaggg	gattctgaaa	cgcagtattg	24420
acataaagga	gtacataaggc	tgaaccaatt	aacctcaagag	tcactcttca	tatggctgac	24480
aatccttgat	gtctggaatt	gatacttacc	catagtaagg	ccatctttat	ctaatttcac	24540
ctcaaatctt	tccacttttg	tataattgcy	atcacctaac	caaccgtcat	ggataaaaagg	24600
aaaatgagac	acgtctaagg	aattatccat	cacacgaaac	gcactagctt	taatcaagta	24660
agacttggtg	taagtcttgt	gataattcgg	atcatcccat	tcaggaaatg	aaggatatatc	24720
attaacagga	tcgcccgaagc	acaccacac	taagccatag	cgctcctggg	agtgatattgt	24780
cctggcttca	gcacttgccg	gtggtaccat	gccagggtga	gctgggatct	gtatgcattt	24840
accagctca	ttgtatctcc	atccgtgata	cggacaaact	aaagtattat	tcgtaatttc	24900
tcccatagac	agaggaacac	ctcgggtggg	gcagtagtca	agccatacct	gtatgggtga	24960
attttgttca	taactgcgcc	ataataccaa	cttactcctc	aacaaacgag	atctggtgat	25020
acttcagggt	ttacagtctt	ctacattggc	gactacgtgc	cagttattga	ttaagattgg	25080
gtcggtagtt	gtcataattg	tctcctagtt	ttgccagcca	gcgaggcgta	agtcagaatt	25140
taagtttatg	cttgtgtttg	agcctgcgat	cgctaaatta	tccttttcaa	ggcatccacc	25200
aacagtgggt	tgatgttgtt	ttttgtaaaa	atcagagtta	gcacctctga	atcggttaatt	25260
gaagtgttgg	cagctgcggg	atgccataca	gttgggtgat	aaaacattgc	tgccctcctt	25320
ggaagtga	gcataatttc	tgcaatttagt	gaattggcag	aagatgaatc	taatgagtgt	25380
tcccattggg	ggctacttgg	tataactcgc	attgtacca	tagtattatc	tgtatcctgt	25440
aagtatatag	ttatgaatac	catggcttga	ttggctactg	gaaccaacaa	ccgaagcgcg	25500
tcgtcattta	actcgttttt	tgacatggat	gcaagtgcgt	tcaatacttc	aactacatat	25560
ccatggtctt	gatgccaagc	aatgtatcct	gtacctgcac	gaattatggc	tagatcggtg	25620
atcaatagga	agatatcaga	cccaattaga	gcctgtactg	gtcccatcac	agttggaagc	25680
tctaaaagcc	tctgaattat	cttttgatac	ctaactggat	ctgggatagt	atgctcagac	25740
caccactcat	agtcaaccgc	caatactccc	ccacgttttt	gttcggtaat	aagttctact	25800
tcatgcgcta	tttcttcaat	taacgctttt	ggtacagctt	cttcaactgt	gaaataacca	25860
tcatttgtgt	aagcttgttt	ttgttccgct	gtgagcatct	ctcttattct	cttgcaattc	25920
aaaggattta	gtggatcgtc	tggacataat	taagggtcaat	actgctgtaa	ctatcaatgg	25980
ttagtaggaa	ttatcctata	gctgttcttt	ctctggatag	aagaaagggt	gtgagaagct	26040
cgctccgact	tcatttcagc	caatttttct	gcagaccaat	actgaaaata	toccaatctt	26100
aataattcat	cactagcctc	ttgttaactgg	ctgaatgact	gtactgatgc	taaaacatac	26160
ttagggtgag	ttatgattac	gttattcaca	ttctccgcgt	catcaccaac	atattgtttg	26220
tctggatgcy	atcctaaagc	taccaaatcg	tattctggta	atacataatt	cgccttggtg	26280
atgtaccttt	ccaacctctg	tgcatctagg	ttttgagggt	cgcagccaaa	aatcaccatt	26340
tcaaagtcac	tattccatgt	tcttatctgt	tccattagaa	gctctggcag	ttcaggtcca	26400
tgaaccaaac	gaacactaac	acggttat	aaccaagctg	ccttcgcgta	aggacagggt	26460
ggaataattc	ctgttagagg	attgggaatg	ctgacaacat	tgataatcca	atcctctatt	26520
tcttggcgaa	attgttccgat	atttatcata	actgttgatt	tttctcctt	tgtagtaatt	26580
agtagttaaa	ggatttagtg	gatattaatc	taggtcatag	tataaccata	tattaggctc	26640

ES 2 525 045 T3

gatgtatatt	cccatattgt	tgggatatgc	aatstttgaca	ggtactaagc	ctttgggaat	26700
aatatagtca	ccagttttctg	gaaaacgcat	cccaactcta	tcttcccaac	cgtcaatagt	26760
atcattaatt	gttgtggatt	taaaacagat	ccctgcaatt	ttagcccat	gtttgacatt	26820
aactcgtaac	caagggtcaa	atataagacc	atsttttatct	cgccaggtaa	tataccgctc	26880
tatgggtata	agtgggtaaa	gatatttttag	gcttggacgt	gcagccatga	tcaaagaatt	26940
aagaccgtgg	tattgagcaa	gttctttcat	gtatccaatc	agatactgac	tcaagttttt	27000
gccttgatac	tctggtagga	ttgaaatcga	tactacacat	aacgcattag	gcaggcgggt	27060
ctgttctcgg	tcttcaagcc	acttggctaa	agcccagtc	caaccttctg	ccggtaaactc	27120
atcaaaacgg	ctttcataag	ttaaagggat	acagtttctt	tgcgctatca	taagctgtgt	27180
ggtagcttct	actaacccaa	actggaattc	tggataaatt	tcaaatagag	ctaaggaagc	27240
tggatctgcc	cagacatcat	gtatcaaaaa	ttttgggtat	gcttgatcaa	agacactcat	27300
cgtcctttcc	acaaaatcag	aagttttctt	tgggggttaca	aagctatact	ctaaattatg	27360
ctgtacaatt	tgaatggtca	tgggttattg	gctaactcct	aaattttatac	tggaaagtcaa	27420
atgagatctc	actatcgtta	ttatctggaa	gtacttgcac	tgtcaattca	ttaccgactt	27480
tccatttccc	aggcataaatt	aataagtttag	ggtagaggtag	aatgccgtcg	tactgtcggg	27540
cgcgggcga	aatgtctgaa	ttctcgccac	catgttttatt	caagaggact	tcaactgggtg	27600
tgatgacaaa	agtcattcct	gacccaaggt	ggcgcgatcg	ccgcttttga	tttgctggag	27660
tggaaacact	aacaaataag	gcacaccctc	ctagagaata	agaccagtta	gcagactgcg	27720
gatcggcaga	ccaatggcag	ggacaagaca	ccgcatcaag	gctatgtaac	gcattcaaaa	27780
aatcaaatgc	ttgacctgca	tattcctcta	ctgttaagaac	tgttggttca	gggtgggaaaa	27840
agatgacaag	tgtcagaaga	tccgcatttt	cgtgctgaag	caattcgttt	tcattaactt	27900
catcaatgta	tttgtagata	ccctcaagcg	tatgctcaac	caagatcggg	tcagttaaag	27960
atgagactat	caggatatcta	atcattccct	tctgttcccc	gatagttccc	cagaagcaag	28020
ggaaggcaga	atcgtgtagt	gtttcaaca	atgttgagta	gctagtgcgt	acccaagcag	28080
gaaggcactc	ctctaagaaga	gaggattcca	tctggctttt	gttccagatt	gggtgaactc	28140
cgtcaggaca	taaattcttg	attaccatag	ctgagttgaa	aagtgagctt	atttatacaa	28200
aaacgatgga	agtgcacact	gatggatggg	acttcaacc	cctacacata	attattatca	28260
ttactatgtg	gcaggtcctt	ctatatctta	ttttttggaa	gtccctgaaa	attattcaac	28320
aagatcgaga	cgttgttggt	gccagaattt	gtgacagcca	ggtcaagctt	gctgtcgccg	28380
ttgaaatccg	caattgctat	agattcagga	ttagtaccga	ctggaaagtt	agtagctatg	28440
ccaaaagacc	cattaccatt	tcttggttaag	accgagacgt	tattgctact	ataatttgta	28500
acagccaggt	caagtttact	gtcgccattc	acatctctaa	tcgctacaga	gtagggatta	28560
gtaccggctg	gaaagttagt	ggctgcgcca	aaagacccat	taccatttcc	cagtaagacc	28620
gagacgttat	tgctgctagt	atlttgcaaca	gccagggtcaa	gcttgctgtc	gccattttaca	28680
tccccagttg	ctacaaatat	gggattagta	ccgactggaa	agttagtggc	tgcgccaaaa	28740
gacccattac	cattttccag	taagaccgag	acgttattgc	tgacccaatt	tgtaatagca	28800
aggtcgagct	tactgtcgct	attaaaatcc	gcaatcgcta	cggaaatcga	ataagtatcg	28860
acaggggaagc	tgctggctgc	gccaaaagac	ccattaccat	ttcccagtaa	aaccaagacc	28920
ttattgtcga	accaatttgt	aaaagcaagg	tcaagctcac	tatcgttatt	cacatctcca	28980
atggctacag	aataagggtt	agtaccaact	gaaaagttag	tggctgcgcc	aaaagaccca	29040
ttaccatttc	ctagtaagac	cgagacgtta	ttgctactaa	aatttgcaac	agccagggtca	29100
agcttgctgt	cgccatttac	atccccagtc	actacaaaga	cgggattagt	accgactgga	29160
aagttagtgg	ctgcgccaaa	agaccatta	ccatttccca	gtaagaccga	gacgttattg	29220
tcgaaccaat	ttgtaacagc	caggctcgagc	ttactatcgc	tattgaaatc	cccaactgct	29280
acagagtcag	catcaagacc	agttgggaag	ttaatagcag	tagcataact	actcctgtgg	29340
gcaaattctca	ctcctacgga	caaatttaacc	ggaacactaa	attgcccaga	aagcttttca	29400
ttcttcagat	aatagtcagt	tatatlttgct	aatgcaacag	gagttatata	taaaaatgta	29460
ctaacagata	atatccccgc	tataattagt	aaagttagcc	ttttcacgag	ttgtatagtt	29520
caaatgtatt	aacaatgttt	gtagccatac	accatcggtg	atgaagaaag	gtattgatcg	29580
caaaatatct	atccttgatc	tagcctatca	cctaagttaa	gccatattga	gttctattta	29640
gattttcttt	ataaatcagc	tataatctat	tgtttgaaaa	ttgtgaattt	gttttccacg	29700
tatttgagta	gttggtctgt	gctttcctcg	acggtgagtt	cggatgtttc	caccataaaa	29760
tctgggctat	tgggtgggttc	ataagggcg	ctgattcccg	taaatccatc	tatttcccca	29820
ctgcgtgctt	ttagataaag	acctttcggg	tcacgctgct	cacaaagtcc	cagtgagggtt	29880
gcaatgtata	cttcatgaaa	tagatctcca	gctagtctac	gcacctgttc	tcggtcattc	29940
ctgtagggtg	agatgaaggc	agtgatcact	aggcatcctg	actccgcaaa	gagtttggga	30000
acctcaccca	aacgcaggat	atlttctgag	cgatcactag	cagaaaatcc	taaatcgga	30060

cacagtccat	gacgaacact	atcacccatct	aaaaaaaagg	tagaccatcc	tttctcgaac	30120
aaagtctgct	ctaatttttaa	agccaatggt	gttttaccag	ccccggacag	tccagtaaac	30180
catagaatcc	cgcttttatg	accattcttt	agataacgat	catatggaga	tataagatgt	30240
tttgtatagt	gaatattagt	tgatttcata	ttgctggagt	ttagactaaa	cagaagagcg	30300
atcgctccat	gcctgagatt	ttagtcagta	ttccactcc	tgtaaacca	caaaaaacac	30360
ggggtaacct	ggaaaattcc	cctggggatc	agctgaaaac	tgctgtttaa	cctgcattat	30420
tcatgaaggc	aaaaacagga	aaaacaaaac	ctaaccattta	tacccaatt	tatggcgga	30480
ctaacttaat	aagtaaaaag	taaattaaac	ctaattaaaa	tccctgattt	taaccccaaa	30540
atcaatattt	taaacctcaa	aacttctctt	aatcccccat	ttagacacac	ctatcctatc	30600
aaggcttaat	tttaagaaaa	aattatttca	aactcgctcg	ccaaacgctc	cataatcaaa	30660
ttattttcag	acgaaaaagg	acagtaatat	ggtagctcta	ccaacaccct	tcttgcgga	30720
actgtcacct	tcgctgctat	tttgataatc	gtttccctta	acctaggaac	ctgggcttta	30780
gccagttttg	ttccctgtgc	tgcttgccga	attcccaaca	ttaaaatgta	agctgcttga	30840
gataaaaaata	accgaaactg	attgacaata	aatttctcac	agctgagctc	atctgatttt	30900
atccccagtt	ttaattccctt	aattctatgc	tctgaagtag	ctcctctttg	aacataaaat	30960
ttatcgtata	aatcctgagc	ttctgtttcc	aagctagtaa	ttataaatct	aggattgggt	31020
cctttttcta	gccattctgc	tttcataatt	actcgccgag	gttctgacca	actccgagct	31080
gcgtaataca	catcatcaaa	taaacgaact	ttttctcctg	tgcgacaata	ttccagctctg	31140
gctcggtaaa	gaaggtaatt	aatttttctg	tttaagacat	cattattgct	gaatccaaaa	31200
acatatccaa	ccccgctttt	ttcacaaaac	tcaatgattt	ctggtaacga	gaaacccccg	31260
tctcccccta	gaacaattct	aatttcaggt	aaggctcttt	tgattcgcaa	aaataaccat	31320
tttagaatgc	cagctactcc	tttaccagag	tgagaatttc	ccgcccttag	ttgtagaact	31380
aatggataac	actgggaagc	ttcattaatc	agaactggaa	agtagatata	atgcctatgg	31440
taaccattaa	ataagctcag	ttgttgatga	ccatgagtta	gagcatocca	cgcatctatg	31500
tccaggacaa	ttcttttttg	ttcccaggga	taggattcta	ggaaatttatc	aacaaataac	31560
cgacgaattt	gtttgatata	tttttgagtc	acctgatttt	ctaaacgact	catagttggt	31620
tgactagcta	ataagttttc	tcctactgtg	ggaacttgat	tacaaactag	cttaaaaaatt	31680
ggactctggc	gcaattttatt	actatcgttg	ctatcttcat	agccagcaat	tatttgataa	31740
attcgttggc	taatttaattg	agaaagagaa	tttttgactt	tagtttggtc	ccgattatcc	31800
gtcaaaacat	ctgccatata	ttgacaaatt	tttacctttt	cttctacttg	tcgtgccaga	31860
ataattccgc	catcactact	taaactcata	tcagaaaaag	tcagatctaa	agttttttta	31920
tcgaagaaat	ttaaagataa	tcttgaggaa	gatttagtca	tatatagtgg	atagggttaa	31980
tttttaaaat	cctgattttat	tatagctggt	tttattcctt	tttttcagtt	tataactaaa	32040
gttagttatt	atttaatttg	gtgacggata	ggaattacag	agtgttgga	tgacaaaatt	32100
gccgtagctg	ttgcagtata	accctttcag	cgatttttat	tctactctga	tgaataatcc	32160
aggataaggct	tgccatcact	ttctgggtag	acaatgtcag	gcgcgattgt	ctccccaccc	32220
tgattaacgt	tagattttat	cacccccagt	tgagtttttg	gtgcaatttc	cctcaccata	32280
tctataacct	ccattcaact	tggtattgac	tcaatcggtt	caatttacta	taacatgact	32340
tatgtggggg	tggtgtgcata	ccctcactta	aaattaatgg	atttgaatct	cctcgactg	32400
ctgcaacttg	aaaaactctg	agagtcagtt	gagagctaac	tctaccagga	ggagagtttt	32460
taaaaacccc	cttcccagagc	gatcgcataa	tttatggtat	acaagaatag	tgggtgaaaa	32520
actaactggc	gatcgctctt	ttcatttaag	agacaccctt	tagttttttt	tgcagctctca	32580
tgaattttaa	cgatatctaa	ttattttcaa	cctatctttg	ccctgtaaca	atgtatgcta	32640
ccctttgacc	aatatttagta	gcatgatctg	ccattctctc	taaacactga	attgctaattg	32700
ttaatagtaa	aatgggctcc	actaccccg	gaacatcttt	ctgctgcgcc	aaattacgat	32760
ataacttttt	gtaagcatca	tctactgtat	catctaataa	tttaatcctt	ctaccactaa	32820
tctcgtctaa	atccgctaaa	gctactagga	tggtagccaa	catagattgg	gcatgatcgg	32880
acataatggc	aacctcccc	aaagtaggat	gggggggata	gggaaatatt	ttcattgcta	32940
tttctgccaa	atctttggca	tagtccccaa	tacgttccaa	gtctctaact	aattgcatga	33000
atgagcttaa	acaccgagat	tcttggtctg	tgggagcttg	actgctcata	attgtggcac	33060
aatcgacttc	tatttgtctg	tagaagcgat	caattttttt	gtctaatactc	cgtatttgct	33120
cagctgctgt	tataatccga	ttgaataggt	cttggtgact	cagacggaat	gactgctcta	33180
ctaaagcacc	catacgcaa	acatctcggt	ccagctcttt	aatggcacgt	ataggttgag	33240
gtttttcaaa	aattgtatat	ttcacacacg	ctttcatatt	tttaatctcg	ggtttaatat	33300
atctttagct	atttatagct	tgattcagaa	atatccgcca	tcagtgtgaa	ccacctgggg	33360
aagatgaatt	tgtatccaag	caccaccggt	atcaggatgg	ttcatggccc	tgattttgcc	33420
accatgagct	ataattattt	ggcggaacat	ggataaccct	aaaccactac	cagtaatttc	33480

tactgtttca	ttctcagagc	gggactcgcg	gtgtctagct	ttgtccccc	gataaaatct	33540
ttgaaagaca	tggggtagat	ccatgggagc	aaatccaacc	ccggaatcaa	taatgttaat	33600
ttctaaaatc	tgattttgata	cttggttttaa	tattgtatct	gcttctggat	caacccccatt	33660
aatagacttc	tccccacaaa	ctggattcat	ttcaatgaaa	atagtaccgt	tcaggttgct	33720
gtattttaata	cagtatatcta	acagattaag	aaacacttga	taaattcttg	acttatcagc	33780
acatatatag	accttttccg	ggccggagta	agaaatacta	agatgctgat	tagcggctag	33840
gggctctaaa	ttctcccaga	ctgaaaaaat	tagggagcgg	acttctagca	tttccaaatt	33900
cagttgtatg	gaggagggtta	tttccatctg	ggtcagggtct	aaccaatttt	ggactaaatt	33960
aattagtctg	tcaacctcct	gcatcaagcg	gatgacccaa	cggtttagag	ggggatctaa	34020
gcgagtttgc	agggtttctg	cgaccagacg	aatggaagtc	agagggtgtc	tcagttcatg	34080
ggccagggtct	gaaaaagagc	ggtcacgttg	ctgatgaatg	tctacaaatt	gttggtgact	34140
ttctagaaac	acaccacttt	gtcccccg	taggggaaaa	ctgtagctg	ctaaagacaa	34200
tggtcttaat	cctaaaatac	cctgaccatg	atctcgggaa	gggtgaaaaa	tccactcttg	34260
cattttcggt	ttttgccaat	cccggttttg	ctcaattaac	tgatccagct	cataggatct	34320
cactaattcc	agtagcaggc	gcacttgacc	cggttgccat	ctttgtaaat	acagcatttc	34380
ccgcgcgcac	tgattacacc	atagtagttg	gttttcttca	tctacttgta	aatatcccaa	34440
aggcgacgca	tccagcaact	gttcataagc	tttgagtgc	aagcgttaagt	tttggtgtct	34500
atctctaacg	gtagatattt	tacgatgtaa	tccagctaatt	aggggttaata	atatcttttc	34560
agcgtgaggg	tttaaggggt	gggttaactg	ctccaaatga	ctgttaagtt	gaaattgttg	34620
ccaaagccaa	aaacccaaac	cgactgccaa	accagaaga	aatcccaata	agaacatttg	34680
atcgtaagt	tgctatttga	ccggaattaa	agggggagga	tccaagcacg	gtctttacag	34740
gacggctttt	tctaattgtt	aaattataat	tataatcgg	agggactgct	ttgggaaaat	34800
gcatcgccc	aggtatctgt	aaccatttct	gtaccacagg	ttagactgga	tcaggtaact	34860
gatacacttc	ttgtctgaat	ttatgtccaa	tcaaaatgac	aactcccaaa	atgataactc	34920
ccgtgacaag	agccaaaac	ccgaatccag	cagatgggtt	aaaataaaaa	gaccacgacc	34980
acctaagga	ataggaaaac	caaaaacaga	atagcccaca	tatagaaatc	aaccaaactc	35040
atagccaaaa	cccctaactg	tgacaatata	ttctggatgg	ctagggtcta	actctaattt	35100
ttccctcagc	catcgaaatg	gaacatccac	cgttttactg	tcaccaacaa	aatcaggacc	35160
ccaaacctgg	tctaataact	gttcccgtga	ccacaccctg	cgagcataac	tcataaatag	35220
ttctagtaac	cggaaattctt	tcgggtgacaa	gctcacctcc	ctccctctca	ctaaccaccg	35280
acattctctga	ggattttaa	tgatatcctt	atatttttaa	gtgggtatca	aggggcaatt	35340
agaaaaccgc	tgacgacgta	acagggcgcg	acacctagcc	accatttccc	gtacgctaaa	35400
aggtcttagt	aggtaatcat	ccgcccctac	ctctaaaccc	agcaccgggt	cagtttctact	35460
acettctcgca	ctcagaattat	aaatcgggtat	gggaattacc	tggtgacgta	acaaacgaca	35520
aatatctaat	ccgttgattt	gtggcaacat	caagtctagc	acaagcaggt	cgaaggataa	35580
ctcaccaggt	tggtgtctcta	aatttctgat	taattccaca	gcacaacgac	catccttagc	35640
agtcacaact	tcataacctt	cacctcttaa	ggctactaca	agcatctctc	ggatcagttc	35700
ttcgtcttcc	actattaaaa	cgcgactaac	tggttcaata	tccgatttag	tgaagtatct	35760
agggttaattc	agtagtattc	attgataaca	aaaatttgta	agaatgtact	ggtctgggtt	35820
tcccactagt	atatgatcct	cactcattga	tgccacatat	tggggaacac	ggaattcttg	35880
tattcaatac	aacaatttgc	ttaaatttat	aattcaaata	ggtgttttat	agaaaatttt	35940
gtcgaatatt	tccacatttg	tggtttttag	ttcaggcaaa	acgagagaag	tctaaagtgg	36000
gtggaatatac	ctgaattctt	ccaggacctc	tgcttctggt	tgcttctggt	aaactaatat	36060
ccccagtata	taggggttta	cccacaatta	ctcctgtaac	cccctgatgt	tctaaagata	36120
ataaggttaa	taggtcagta	acagaaccca	caccccaga	ggcaatcacg	ggtatggaaa	36180
tagcagatac	caagtctctt	aatgctcgca	agtttggtcc	ctgaagcgta	ccatcacgggt	36240
ttatatccgt	ataaataata	gctgccgcac	ccaattctctg	catttggtt	gctagttggg	36300
ggggcaaaat	ttgagaatct	tctaaccaac	ccctggtagc	aactagacca	ttccgcgcac	36360
caatcccaat	tataatttgc	tggggggaatt	gttcacacag	tccttgaaac	agatctggtt	36420
gctctactgc	tacagttccc	agaattgccc	actgtacccc	aagattaaat	aactgtataa	36480
cgctggagct	atcacgtatt	cctccgccaa	cttcaatagg	tatggaaata	gcattggtaa	36540
tagcttctat	agtagataaa	ttactattt	taccagtttt	tgctccatct	aaatctacta	36600
aatgtagtct	tggtgtctct	tggtctgccc	acatttttagc	ggtttccaca	gggttatggc	36660
tgtaaacctg	ggatttgtgca	tagtcacctt	tgtagagtct	tacacaacgc	ccctctaata	36720
gatctattgc	tgggataact	tccatgacta	attagtgaat	agggttaattt	cagttgagct	36780
aaatggagaa	ggaggggattc	gaacctcgg	atggacctta	cgattccatc	aacagatttag	36840
caatctgcgc	ctttcgacca	ctcagccacc	tctccaggtt	tggtataaat	tatgatgggt	36900
caatcctaac	agacaatttt	tggtctgtca	agagattttt	tgcaagtgga	ggaggaaatc	36960
cgtcagggat	ttcaatcctg	gtcaactttt	ttttgatatt	gaatataaag	ttaagtttaa	37020
caatttcttag	tggcgctcct	ccaacagtag	atataaaata	tgagttgggtc	cacaatgaag	37080
gacgtcttga	ttttaatagt	caaatccctc	caaattccatt	ataatcccat	gaatgctctt	37140
tcaattccta	cctggattat	ccatatttct	agtgctcattg	aatgggtagt	tgccatttcc	37200
ctcatctgga	aatatggcga	actgacccaa	aaccatagtt	ggaggggatt	tgcccttaggt	37260
atgatacccg	ccttaatttag	cgccctatcc	gcttgtacct	ggcattattt	cgataatccc	37320
cagtccttag	aatgggttagt	cacctccag	gctactacta	cgttaatagg	taattttact	37380
ctttgggcag	cagcagctctg	ggtttggcgt	tctactcgac	cgaatgaggt	tctcagtatc	37440
tcaataaag	agtagaccgt	tatgatgtca	aaagaaactc	tctttgctct	ctccctgttc	37500
ccctatttgg	gaatgttgtg	gtttctcagt	cgcagtcctc	aaatgcccc	ttaagggtctc	37560
tatggattct	atggcacttt	agtatttgtt	ggtgttacca	ttccag		37606

<210> 2
 <211> 1320
 <212> ADN
 <213> *Cylindrospermopsis raciborskii* T3

5

<400> 2

```

atgatcccag ctaaaaaagt ttatttttta ttgagtttag caatagttat ttcacccttt    60
ttatccatga ttgtgggtat ttacgaaaat attaaattta ggggtattatt tgatttggtg    120
gtcagggcac taatgggtgt tgactgcttc aatatcaaaa aacatcgggt caaaattagt    180
cgtcaattac ctctacgttt atctattgga cgtgagaatt tagtaatatt gaaggtagag    240
tctgggaatg tcaatagtgc tattcaaatt cgtgattact atcccacaga atttcccgta    300
tccacatcta acctgatagt taaccttccc cctaatacata cttaggaagt aaagtacacc    360
attcgacctt atcaacgggg agaatttttg tggggaaata ttcaagttcg acagctggga    420
aattggtctc taggggtggga caattggcaa attcccaaaa aaactgtggc taagggtgat    480
cctgatttgt taggactcag atccctcgct attcgtttta ccctacaatc ttctggatct    540
atcactaaat tgcgtcaacg ggggaatggga acggaatttg ccgaactccg taattactgc    600
atgggggatg atctacggtt aattgattgg aaagctacag cttagacgtgc ttatggaaat    660
ctgagtcctc tagtaagagt tttagagcct caacagggaac aaactctgct tatattatta    720
gatcgtggta gactaatgac agctaattgta caagggttaa aacgatatga ttgggggtta    780
aataccacct tgtcttttgg atttagcagga ttacataggg gcgatcgcgt aggagtaggg    840
gtatttgact ccagactgca tacctggata cctccagagc gaggacaaaa tcatctcaat    900
cggcttatag acagacttac acctattgaa ccagtgttag tggagtctga ttatttaaatt    960
gccattacct atgtagtaaa acaacagact cgtagatctc tagtagtggt aattactgat   1020
ttagtctgat ttactgcttc ccatgaacta ctagttagcgc tgtgtaaatt agtgcctcga   1080
tatctacctt tttgtgtaac actcagggat cctgggattg ataaaatagc tcataatttt   1140
agtcaagact taacacaggc ttataatcga gcagtttctt tggacttgat atcacaaaga   1200
gaaattgctt ttgctcagtt gaaacaacag ggagtttttg tgttggtatg accagcaaatt   1260
caaatttccg agcagttggt agaaaggtac ttacaaatca aagccaaaaa tcagatttga   1320
    
```

10 <210> 3
 <211> 439
 <212> PRT
 <213> *Cylindrospermopsis raciborskii* T3

15 <400> 3

```

Met Ile Pro Ala Lys Lys Val Tyr Phe Leu Leu Ser Leu Ala Ile Val
1           5           10           15
Ile Ser Pro Phe Leu Ser Met Ile Val Gly Ile Tyr Glu Asn Ile Lys
           20           25           30
Phe Arg Val Leu Phe Asp Leu Val Arg Ala Leu Met Val Val Asp
           35           40           45
Cys Phe Asn Ile Lys Lys His Arg Val Lys Ile Ser Arg Gln Leu Pro
           50           55           60
Leu Arg Leu Ser Ile Gly Arg Glu Asn Leu Val Ile Leu Lys Val Glu
65           70           75           80
Ser Gly Asn Val Asn Ser Ala Ile Gln Ile Arg Asp Tyr Tyr Pro Thr
    
```

				85					90				95		
Glu	Phe	Pro	Val	Ser	Thr	Ser	Asn	Leu	Ile	Val	Asn	Leu	Pro	Pro	Asn
			100					105					110		
His	Thr	Gln	Glu	Val	Lys	Tyr	Thr	Ile	Arg	Pro	Asn	Gln	Arg	Gly	Glu
		115					120					125			
Phe	Trp	Trp	Gly	Asn	Ile	Gln	Val	Arg	Gln	Leu	Gly	Asn	Trp	Ser	Leu
	130				135					140					
Gly	Trp	Asp	Asn	Trp	Gln	Ile	Pro	Gln	Lys	Thr	Val	Ala	Lys	Val	Tyr
145				150					155					160	
Pro	Asp	Leu	Leu	Gly	Leu	Arg	Ser	Leu	Ala	Ile	Arg	Leu	Thr	Leu	Gln
			165					170					175		
Ser	Ser	Gly	Ser	Ile	Thr	Lys	Leu	Arg	Gln	Arg	Gly	Met	Gly	Thr	Glu
			180					185					190		
Phe	Ala	Glu	Leu	Arg	Asn	Tyr	Cys	Met	Gly	Asp	Asp	Leu	Arg	Leu	Ile
	195						200				205				
Asp	Trp	Lys	Ala	Thr	Ala	Arg	Arg	Ala	Tyr	Gly	Asn	Leu	Ser	Pro	Leu
	210				215						220				
Val	Arg	Val	Leu	Glu	Pro	Gln	Gln	Glu	Gln	Thr	Leu	Leu	Ile	Leu	Leu
225					230					235				240	
Asp	Arg	Gly	Arg	Leu	Met	Thr	Ala	Asn	Val	Gln	Gly	Leu	Lys	Arg	Tyr
			245					250					255		
Asp	Trp	Gly	Leu	Asn	Thr	Thr	Leu	Ser	Leu	Ala	Leu	Ala	Gly	Leu	His
			260					265					270		
Arg	Gly	Asp	Arg	Val	Gly	Val	Gly	Val	Phe	Asp	Ser	Gln	Leu	His	Thr
	275						280					285			
Trp	Ile	Pro	Pro	Glu	Arg	Gly	Gln	Asn	His	Leu	Asn	Arg	Leu	Ile	Asp
	290					295					300				
Arg	Leu	Thr	Pro	Ile	Glu	Pro	Val	Leu	Val	Glu	Ser	Asp	Tyr	Leu	Asn
305				310						315				320	
Ala	Ile	Thr	Tyr	Val	Val	Lys	Gln	Gln	Thr	Arg	Arg	Ser	Leu	Val	Val
			325						330				335		
Leu	Ile	Thr	Asp	Leu	Val	Asp	Val	Thr	Ala	Ser	His	Glu	Leu	Leu	Val
		340					345					350			
Ala	Leu	Cys	Lys	Leu	Val	Pro	Arg	Tyr	Leu	Pro	Phe	Cys	Val	Thr	Leu
	355					360					365				
Arg	Asp	Pro	Gly	Ile	Asp	Lys	Ile	Ala	His	Asn	Phe	Ser	Gln	Asp	Leu
	370				375						380				
Thr	Gln	Ala	Tyr	Asn	Arg	Ala	Val	Ser	Leu	Asp	Leu	Ile	Ser	Gln	Arg
385				390					395					400	
Glu	Ile	Ala	Phe	Ala	Gln	Leu	Lys	Gln	Gln	Gly	Val	Leu	Val	Leu	Asp
			405					410					415		
Ala	Pro	Ala	Asn	Gln	Ile	Ser	Glu	Gln	Leu	Val	Glu	Arg	Tyr	Leu	Gln
		420					425					430			
Ile	Lys	Ala	Lys	Asn	Gln	Ile									
		435													

<210> 4
 <211> 759
 5 <212> ADN
 <213> *Cylindrospermopsis raciborskii* T3

<400> 4

	atgatagata	caatatcagt	actattaaga	gagtggactg	taatttcctt	tacagggtta	60
	gccttctggc	tttgggaaat	tcgctctccc	ttccatcaaa	ttgaatacaa	agctaaattc	120
	ttcaaggaat	tgggatgggc	gggaatatca	ttcgtcttta	gaaatgttta	tgcatatgtt	180
	tctgtggcaa	ttataaaaact	attgagttct	ctatttatgg	gagagtcagc	aaattttgca	240
	ggagtaatgt	atgtgccctt	ctggctgagg	atcatcactg	catatatatt	acaggactta	300
10	actgactatc	tattacacag	gacaatgcat	agtaatcagt	ttctttggtt	gacgcacaaa	360
	tggtcatcatt	caacaaagca	atcatggttg	ctgagtggaa	acaaagatag	ctttaccggc	420
	ggacttttat	atactgttac	agctttgttg	tttccactgc	tggaatttcc	ctcagagggt	480
	atgtctgtag	tggcagtaga	tcaagtgttg	cataacaatt	ggatacacct	caatgtaaag	540
	tggaaactcct	ggtaggaat	aattgaatgg	atttatgtta	cgccccgtat	tcacactttg	600
	catcatcttg	atacaggggg	aagaaatttg	agttctatgt	ttactttcat	cgaccgatta	660
	tttggaaacct	atgtgtttcc	agaaaacttt	gatatagaaa	aatctaaaaa	tagattggat	720
	gatcaatcag	taacgggtgaa	gacaattttg	ggtttttaa			759

<210> 5
 15 <211> 252
 <212> PRT

<213> Cylindrospermopsis raciborskii T3

<400> 5

```

Met Ile Asp Thr Ile Ser Val Leu Leu Arg Glu Trp Thr Val Ile Ser
1      5      10      15
Leu Thr Gly Leu Ala Phe Trp Leu Trp Glu Ile Arg Ser Pro Phe His
20     25     30
Gln Ile Glu Tyr Lys Ala Lys Phe Phe Lys Glu Leu Gly Trp Ala Gly
35     40     45
Ile Ser Phe Val Phe Arg Asn Val Tyr Ala Tyr Val Ser Val Ala Ile
50     55     60
Ile Lys Leu Leu Ser Ser Leu Phe Met Gly Glu Ser Ala Asn Phe Ala
65     70     75     80
Gly Val Met Tyr Val Pro Leu Trp Leu Arg Ile Ile Thr Ala Tyr Ile
85     90     95
Leu Gln Asp Leu Thr Asp Tyr Leu Leu His Arg Thr Met His Ser Asn
100    105    110
Gln Phe Leu Trp Leu Thr His Lys Trp His His Ser Thr Lys Gln Ser
115    120    125
Trp Trp Leu Ser Gly Asn Lys Asp Ser Phe Thr Gly Gly Leu Leu Tyr
130    135    140
Thr Val Thr Ala Leu Trp Phe Pro Leu Leu Asp Ile Pro Ser Glu Val
145    150    155    160
Met Ser Val Val Ala Val His Gln Val Ile His Asn Asn Trp Ile His
165    170    175
Leu Asn Val Lys Trp Asn Ser Trp Leu Gly Ile Ile Glu Trp Ile Tyr
180    185    190
Val Thr Pro Arg Ile His Thr Leu His His Leu Asp Thr Gly Gly Arg
195    200    205
Asn Leu Ser Ser Met Phe Thr Phe Ile Asp Arg Leu Phe Gly Thr Tyr
210    215    220
Val Phe Pro Glu Asn Phe Asp Ile Glu Lys Ser Lys Asn Arg Leu Asp
225    230    235    240
Asp Gln Ser Val Thr Val Lys Thr Ile Leu Gly Phe
245    250

```

<210> 6

<211> 396

<212> ADN

<213> Cylindrospermopsis raciborskii T3

<400> 6

```

tcacccccaa cttattgcag aaaaactttt ttctcttagg taataaatta gtagtttaat    60
tgaaaagcat agcatctctt ttgacttgga ataacaaaat gtcttacgat gtagtctagc    120
taaatagtga cgcaaacgac tgttttctcc ctcaactcta gtcattgatg ttttactaat    180
aatttggtct ccatcgggaa taaatttttg gtaaacttta tagccatccg taatccaaaa    240
ataggatttc caatgctcta tctttttcca taatttgga aatgttttg cacttctatc    300
tcccactaca tattgaataa ttcccgaacg ttgttatct acaactgtcc agaccatata    360

cttggtttttt ttaccaata aatgtttcca actcat                                396

```

<210> 7

<211> 131

<212> PRT

<213> Cylindrospermopsis raciborskii T3

<400> 7

ES 2 525 045 T3

Met Ser Trp Lys His Leu Leu Val Lys Lys Asn Lys Ile Trp Val Trp
 1 5 10 15
 Thr Val Val Asp Asn Lys Arg Ser Gly Ile Ile Gln Tyr Val Val Gly
 20 25 30
 Asp Arg Ser Ala Lys Thr Phe Ala Lys Leu Trp Lys Lys Ile Glu His
 35 40 45
 Trp Lys Ser Tyr Phe Trp Ile Thr Asp Gly Tyr Lys Val Tyr Pro Lys
 50 55 60
 Phe Ile Pro Asp Gly Asp Gln Ile Ile Ser Lys Thr Ser Met Thr Arg
 65 70 75 80
 Val Glu Gly Glu Asn Ser Arg Leu Arg His Tyr Leu Ala Arg Leu His
 85 90 95
 Arg Lys Thr Phe Cys Tyr Ser Lys Ser Lys Glu Met Leu Cys Phe Ser
 100 105 110
 Ile Lys Leu Leu Ile Tyr Tyr Leu Arg Glu Lys Ser Phe Ser Ala Ile
 115 120 125
 Ser Trp Gly
 130

<210> 8

<211> 360

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 8

ttatctacaa ctgtccagac ccatatcttg ttttttttta ccaataaatg tttccaactc 60
 atccagttga caaacttcag gtgtttggga attattatta ctatctgata actgacgacc 120
 tagctttttg acccaacgaa tgactgtatt gtgatttact ttagtcattc tttcaattgc 180
 cctaaatcca ttcccattta catacatggt taaacatgct tcctttactt cttgggaata 240
 acctctagga gaataagatt caataaattg acgaccacaa ttcttgcat gataattttg 300
 10 ttttcccctt ctctggccat tttttctaatt attattggaa tcacagtttg aacagttcat 360

<210> 9

<211> 119

<212> PRT

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 9

Met Asn Cys Ser Asn Cys Asp Ser Asn Asn Ile Arg Lys Asn Gly Gln
 1 5 10 15
 Arg Arg Gly Lys Gln Asn Tyr Gln Cys Lys Asn Cys Gly Arg Gln Phe
 20 25 30
 Ile Glu Ser Tyr Ser Pro Arg Gly Tyr Ser Gln Glu Val Lys Glu Ala
 35 40 45
 Cys Leu Thr Met Tyr Val Asn Gly Asn Gly Phe Arg Ala Ile Glu Arg
 50 55 60
 Met Thr Lys Val Asn His Asn Thr Val Ile Arg Trp Val Lys Lys Leu
 65 70 75 80
 Gly Arg Gln Leu Ser Asp Ser Asn Asn Asn Ser Gln Thr Pro Glu Val
 85 90 95
 Cys Gln Leu Asp Glu Leu Glu Thr Phe Ile Gly Lys Lys Lys Gln Asp
 100 105 110
 20 Met Gly Leu Asp Ser Cys Arg
 115

<210> 10

<211> 354

25 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 10

ES 2 525 045 T3

```

ttatgacctc attttcattt ctagacgttc agcaacgggc attaactcac gtatcagatc      60
aaagtttcct acgttccgtc tcatccagtc taataagaat ttttctcctt catctagctt      120
acctttatca tcaacaaaaa ccatctgctc gcaccaatct acaaatccgg aattagtcac      180
ctcatagact aaaatgatgg gaggaaagtg tgcgaatccc attttttcaa tgacttccat      240
acaaaccagc ttaaataactt gttcgtttgt caattcatta gacataaaga attttccttt      300
aatcaattct gtttctaate ctaccacaga gtaataactc ttggtctgga acat          354

```

<210> 11

<211> 117

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 11

```

Met Phe Gln Thr Lys Ser Tyr Tyr Ser Val Val Gly Leu Glu Thr Glu
1      5      10      15
Leu Ile Lys Gly Lys Phe Phe Met Ser Asn Glu Leu Thr Asn Glu Gln
20     25     30
Val Phe Lys Leu Val Cys Met Glu Val Ile Glu Lys Met Gly Phe Ala
35     40     45
His Phe Pro Pro Ile Ile Leu Val Tyr Glu Met Thr Asn Ser Gly Phe
50     55     60
Val Asp Trp Cys Glu Gln Met Val Phe Val Asp Asp Lys Gly Lys Leu
65     70     75     80
Asp Glu Gly Glu Lys Phe Leu Leu Asp Trp Met Arg Arg Asn Val Gly
85     90     95
Asn Phe Asp Leu Ile Arg Glu Leu Met Pro Val Ala Glu Arg Leu Glu
100    105    110
Met Lys Met Arg Ser
115

```

10

<210> 12

<211> 957

<212> ADN

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 12

```

tcataactta ttacttgacg gagttgcagg ggcataacctt aacttgacct tgggagcgat      60
agaagaaaagg aaggcttcag tgacgggtct ttgactaatc ccagtttcca cttcaactaa      120
aacagcatca caaatgtcga atagtgattg agaatatcta ttcatattca tgaaagtcag      180
agcagattcc atcggagaca tggatgaatt aaaggcagcg ttttcagcgt atcgacctgt      240
aaatatattc ccgtgggaat cttttaacgc taccctgca aaatttttcg tgtagggagc      300
ataactttga ttggcagcgg atagagcagc aagcacaaca tcatcggtag aataggtctc      360
cagatcatga aatactgttt gcattaatcc acctgtgagt cctagatccg ctggtccaaa      420
tggctcgggt agaaaatgtg ggagtttatt tgaggataaa gtttgctcag gctgtgattc      480
attagacttc acaagaagaa caaaattttg atttacagtt gccatctcgt ataaaaattg      540
tcggcagtat ccacatggtg cttcgtggat tgctaatgct tgtaaaccgg tttctccgtg      600
caaccacgca tttatggtgg cggattgttc tgcgtgaact gagaaactaa gtgcctgtcc      660
tacaaattcc atgtcggcac caaataaaag agttccagaa cccagttgat tcttagattg      720
tggtttacca agagcgatcg cccctacata aaactgcgat attggtaccc tagcataagt      780
tgcggtacag ggtagtaatt gaatcattaa cgtactaata ttagtaccaa gtcgatcaat      840
ccaagatgcg acaacacttg agtcaattac agcatgttgg gcaagaattg tccttaactc      900

```

20

tgattgaatg gaacgtggaa ccttggcaat cgctgttct aatgctacat gggtcat 957

<210> 13

<211> 318

25 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 13

Met Thr His Val Ala Leu Glu Gln Ala Ile Ala Lys Val Pro Arg Ser
1 5 10 15
Ile Gln Ser Glu Leu Arg Thr Ile Leu Ala Gln His Ala Val Ile Asp
20 25 30
Ser Ser Val Ala Ser Trp Ile Asp Arg Leu Gly Thr Asn Ile Ser
35 40 45
Thr Leu Met Ile Gln Leu Leu Pro Val Ala Ala Thr Tyr Ala Arg Val
50 55 60
Pro Ile Ser Gln Phe Tyr Val Gly Ala Ile Ala Leu Gly Lys Pro Gln
65 70 75 80
Ser Lys Asn Gln Leu Gly Ser Gly Thr Leu Tyr Phe Gly Ala Asp Met
85 90 95
Glu Phe Val Gly Gln Ala Leu Ser Phe Ser Val His Ala Glu Gln Ser
100 105 110
Ala Thr Ile Asn Ala Trp Leu His Gly Glu Thr Gly Leu Gln Ala Leu
115 120 125
Ala Ile His Glu Ala Pro Cys Gly Tyr Cys Arg Gln Phe Leu Tyr Glu
130 135 140
Met Ala Thr Val Asn Gln Asn Phe Val Leu Leu Val Lys Ser Asn Glu
145 150 155 160
Ser Gln Pro Glu Gln Thr Tyr Thr Ser Asn Lys Leu Pro His Phe Leu
165 170 175
Pro Glu Pro Phe Gly Pro Ala Asp Leu Gly Leu Thr Gly Gly Leu Met
180 185 190
Gln Thr Val Phe His Asp Leu Glu Thr Tyr Ser Thr Asp Asp Val Val
195 200 205
Leu Ala Ala Leu Ser Ala Ala Asn Gln Ser Tyr Ala Pro Tyr Thr Lys
210 215 220
Asn Phe Ala Gly Val Ala Leu Lys Asp Ser His Gly Asn Ile Phe Thr
225 230 235 240
Gly Arg Tyr Ala Glu Asn Ala Ala Phe Asn Ser Ser Met Ser Pro Met
245 250 255
Glu Ser Ala Leu Thr Phe Met Asn Met Asn Arg Tyr Ser Gln Ser Leu
260 265 270
Phe Asp Ile Cys Asp Ala Val Leu Val Glu Val Glu Thr Gly Ile Ser
275 280 285
Gln Arg Pro Val Thr Glu Ala Phe Leu Ser Ser Ile Ala Pro Lys Val
290 295 300
Lys Leu Arg Tyr Ala Pro Ala Thr Pro Ser Ser Asn Lys Leu
305 310 315

<210> 14

<211> 3738

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 14

10 ttaatgcttg agtatgtttt cctcctgctt acaaggcaaa gctttccttt tttgtagcaa 60
atcccaaact gctttgagag atttaattgc ttggtctatc tcctcttcgg tattggcggc 120
tgtaatcgaa aaccttaaag cacttttatt taaaggtagc attggaaaaa tagcaggagt 180
aattaaaata ccatattccc aaaggagtgg acacacatca atcatgtgtt gagcatctcc 240

cactaacacg	cctacgatgg	gaacgtaacc	atagttatcc	acttcgaatc	caatggctct	300
tgcttggtgta	accaaatttgt	gagttagggtg	ataaatttgt	tttcttaact	gctccccctc	360
ctgacgattc	acctgtaatc	cggctaaggc	acttgccaaa	ctcgcaacag	gagaaggacc	420
agaaaatatg	gcagtcacaag	cggtgcgga	gttggttttg	atccggcgat	cgccacaagt	480
taagaatgct	gcgtaagaag	aataggcttt	ggacaaacca	gctacataga	tgatattatc	540
ctctgcaaac	cgcagggtcaa	aataattcac	catcccgttt	cctttgtaac	cgtaaggcat	600
atcgctgctg	ggatttttcg	ccaaaatgcc	aaaaccatga	gcatcatcca	tgtaaattaa	660
ggcattgtac	tcttttgcca	gatgcacgta	agctggcaga	tcgggaaaat	ctgccgacat	720
ggaatacacg	ccatcaatga	caataatctt	tacttggtca	ggcggtatatt	ttgctagttt	780
ttcggctaaa	tcgttcaaat	cattatgtcg	atattggatg	aactgggctc	ctttgtgctg	840
agccagacag	cacgcttcac	aaatacaacg	atgtgcagct	atgtcaccaa	agatgacacc	900
attattccca	gttaatatgtg	gtaaaattcc	tatctgaagc	agtgttacag	ctggaaatac	960
taaaacatca	ggtacgccta	aaagtttgga	caattcttcc	tccaattcct	cataaattgc	1020
tggggaaagca	acaagccgag	tccagcttgg	atgtgtgccc	catttatcca	aagctggtgg	1080
aattgcttcc	ttaacttttg	gatgcaagtc	aagacctaaa	tagttgcaag	aagcaaaagtc	1140
tatcacccaa	tgtcctgcaa	ttagcacctt	gcgaccttgt	tgttctgtga	cgactcttgt	1200
gacttgagga	attttttgtt	ggttaactac	gttttccaga	gtgttgattt	cgttggctga	1260
gtcaacaggt	ggagctagat	cagattgttt	ctcttgtagc	acttggtttt	ggaaataagt	1320
gatgatggca	gttgagagtgt	tcttttgtaa	aaagaacggt	ccagacagat	tgatccctaa	1380
acgttctctc	aggagcggtt	gcagttctaa	taaatctaaa	gaatctaate	ccatatccag	1440
cagtttttgt	tgtggagcgt	aggctgcctg	acgttgggaa	cccattactt	ttaagatgca	1500
ttctttaacg	agatccgcta	cagttttgtt	ttccttagtt	gcagatgttg	cttttggtac	1560
caatgaacca	attgctgagt	taataacggt	tcctttgcga	tcaccaggcg	agtgc aaagc	1620
actgtcgcgc	aggttatatt	caatcaaaat	acccatgccc	agattatctg	tatcttccgg	1680
acgataatta	gcaataattc	ccctaatttc	ggctcctccc	gacacatgga	aaccacaaat	1740
tggtaccaga	agctgtcgtt	gctcattgtg	tagctttaaa	tactccatca	tcggcatttg	1800
ggaataattg	acataatttc	gacagcgagt	tacacccacc	acgctctcaa	tgccgccttt	1860
cagggtagacg	tagtaaagca	taaagtccc	caattcattt	cctaaccccc	gcgcctgaaa	1920
ctcaggtaga	atatttagtg	cgagcagttg	aataactgac	ccttggggag	tatgtaacgt	1980
cggcacttgc	gcataattta	cattctctaa	tgctcagtg	ctggtaattg	tttggaata	2040
aatcgaccca	ataatttgat	cttctataat	cagcactaaa	ttaccttgcg	ggtttagctc	2100
aagtctctgc	cgaatttcat	gagtagatgc	ccgtaaat	tctggccaac	acttgacctc	2160
caagtcaact	aaggcaggta	aatctgacaa	ataggcatga	ctaattttgt	aaggctcttt	2220
ctcgaagtaa	ttaagcgtaa	tgcgagtaaa	aggaaatgtt	tttgggtatc	ttttagaaag	2280
ctctagtttt	ggaaatagac	ctacttgtgc	agcagacatg	agaaaaacct	cagcttccac	2340
aagataactgc	tgagaaaatc	cctgaaacgc	atcgaaatgt	aagtttttcg	ttttgtctaa	2400
aaactgatag	actacccttg	gttccaaaca	atggacctcc	aaaatcatta	aaccgtgttt	2460
attgaccact	tgagaccatc	tttctaagtg	ttccaccaa	ctttgcacca	taacatgagg	2520
aggaataaagc	tctccttgat	catcgacaca	gactgattgg	taaggtaagt	gagcacgttc	2580
tttcaattcg	ttctctttct	gaggaggaat	aaagagacga	tcatggctga	ggaacgaacg	2640
gatgtgcagg	atattttcgg	gatcatgaat	gccatgagct	tctaaagaac	gcaccatttg	2700
ttctgggttc	ccaatatctc	cctgtaaaac	taagtgggga	aggctagcaa	gggtgcgtgt	2760
ggtagctttt	aaagaagctt	cggtataatc	tacacctata	agacgcaggg	gatactgttc	2820
gagtgtcttt	cccttagcag	acttaaatgt	aatggtttcc	cagactcggt	tcaggagagt	2880
tccatcgcca	caccccatgt	cagtaatgta	tttgggttgt	tcttctaattg	gcaactgatt	2940
gaatactgag	agataacttt	cttctaatac	ggcaaaatat	ttctgggtgt	gaaatccact	3000
cccgatcacg	ttaaggggtc	gatcaatgtg	cctttcgtga	ccggaagcat	ctctttggaa	3060
tacggagaga	caattgcca	acaatacatc	atgaatgcgg	gacaacatag	gagtgtagga	3120
cgccactatg	gctgtattca	aggctcgctc	tcccataaat	cgaccaagtt	cggttatggg	3180
caaacgacct	gctgtaagg	cagcccagcc	aagggtggga	aataacttac	ccaactcttc	3240
ttgcactggt	gagcttaagt	aggagagcaa	aggtttgc	tccgaatctg	caagcaagtt	3300
gtgtttgtgc	agtgcagca	ggagtgggat	gaccagtaat	ccatctaaaa	aatctgccat	3360
taggggattg	tccaggttcc	acaattggca	agaacgctca	atccatcttc	ccagcaaatt	3420
tccttggttc	ccttctaata	aagactgaat	tggtagggtg	tacaattgaa	gaatgtcttc	3480
cgaaattttg	ttgtgaatcg	ctgcttctgc	ggtagagag	tatttaagct	ccttatttgc	3540
ggaaagccaa	tgtaaaagct	cgagcatcct	caaagcaact	tgaaatgtc	cgctgttagc	3600
tcccagatgt	tccaccattt	ggtttaaaga	gagaggactt	tcatcggcga	gtaattcaaa	3660

aacacctttt	tctcgacacg	caagaataac	gggaaccgcc	acaaagccgt	gagtataacg	3720
attaatcttt	tgtaacat					3738

- 5 <210> 15
 <211> 1245
 <212> PRT
 <213> *Cylindrospermopsis raciborskii* T3

- 10 <400> 15

Met	Leu	Gln	Lys	Ile	Asn	Arg	Tyr	Thr	His	Gly	Phe	Val	Ala	Val	Pro
1				5					10					15	
Val	Ile	Leu	Ala	Cys	Arg	Glu	Lys	Gly	Val	Phe	Glu	Leu	Leu	Ala	Asp
			20					25					30		
Glu	Ser	Pro	Leu	Ser	Leu	Asn	Gln	Met	Val	Glu	His	Leu	Gly	Ala	Asn
		35					40					45			
Ser	Gly	His	Phe	Gln	Val	Ala	Leu	Arg	Met	Leu	Glu	Ser	Leu	His	Trp
	50					55				60					
Leu	Ser	Arg	Asn	Lys	Glu	Leu	Lys	Tyr	Ser	Leu	Thr	Ala	Glu	Ala	Ala
65					70					75					80
Ile	His	Asn	Lys	Ile	Ser	Glu	Asp	Ile	Leu	Gln	Leu	Tyr	Asn	Leu	Pro
			85						90					95	
Ile	Gln	Ser	Tyr	Leu	Glu	Gly	Lys	Gln	Gly	Asn	Leu	Leu	Gly	Arg	Trp
			100					105					110		
Ile	Glu	Arg	Ser	Cys	Gln	Leu	Trp	Asn	Leu	Asp	Asn	Pro	Leu	Met	Ala
		115					120					125			
Asp	Phe	Leu	Asp	Gly	Leu	Leu	Val	Ile	Pro	Leu	Leu	Leu	Ala	Leu	His
	130					135				140					
Lys	His	Asn	Leu	Leu	Ala	Asp	Ser	Glu	Asp	Lys	Pro	Leu	Leu	Ser	Ser
145					150					155					160
Leu	Ser	Ser	Thr	Val	Gln	Glu	Glu	Leu	Gly	Lys	Leu	Phe	Leu	His	Leu
			165						170					175	
Gly	Trp	Ala	Asp	Leu	Thr	Ala	Gly	Arg	Leu	Thr	Ile	Thr	Glu	Leu	Gly
			180					185					190		
Arg	Phe	Met	Gly	Glu	Arg	Ala	Leu	Asn	Thr	Ala	Ile	Val	Ala	Ser	Tyr
		195					200					205			
Thr	Pro	Met	Leu	Ser	Arg	Ile	His	Asp	Val	Leu	Phe	Gly	Asn	Cys	Leu
	210					215					220				
Ser	Val	Phe	Gln	Arg	Asp	Ala	Ser	Gly	His	Glu	Arg	His	Ile	Asp	Arg
225					230					235					240
Thr	Leu	Asn	Val	Ile	Gly	Ser	Gly	Phe	Gln	His	Gln	Lys	Tyr	Phe	Ala
			245						250					255	
Asp	Leu	Glu	Glu	Ser	Ile	Leu	Ser	Val	Phe	Asn	Gln	Leu	Pro	Leu	Glu
			260					265					270		
Glu	Gln	Pro	Lys	Tyr	Ile	Thr	Asp	Met	Gly	Cys	Gly	Asp	Gly	Thr	Leu
		275					280					285			
Leu	Lys	Arg	Val	Trp	Glu	Thr	Ile	Gln	Phe	Lys	Ser	Ala	Arg	Gly	Lys
	290					295					300				
Ala	Leu	Glu	Gln	Tyr	Pro	Leu	Arg	Leu	Ile	Gly	Val	Asp	Tyr	Asn	Glu
305					310					315					320
Ala	Ser	Leu	Lys	Ala	Thr	Thr	Arg	Thr	Leu	Ala	Ser	Leu	Pro	His	Leu
			325						330					335	
Val	Leu	Gln	Gly	Asp	Ile	Gly	Asn	Pro	Glu	Gln	Met	Val	Arg	Ser	Leu
			340					345					350		
Glu	Ala	His	Gly	Ile	His	Asp	Pro	Glu	Asn	Ile	Leu	His	Ile	Arg	Ser
		355					360					365			
Phe	Leu	Asp	His	Asp	Arg	Leu	Phe	Ile	Pro	Pro	Gln	Lys	Arg	Asn	Glu
370						375					380				

Leu	Lys	Glu	Arg	Ala	His	Leu	Pro	Tyr	Gln	Ser	Val	Cys	Val	Asp	Asp
385					390					395					400
Gln	Gly	Glu	Leu	Ile	Pro	Pro	His	Val	Met	Val	Gln	Ser	Leu	Val	Glu
				405					410						415
His	Leu	Glu	Arg	Trp	Ser	Gln	Val	Val	Asn	Lys	His	Gly	Leu	Met	Ile
			420					425					430		
Leu	Glu	Val	His	Cys	Leu	Glu	Pro	Arg	Val	Val	Tyr	Gln	Phe	Leu	Asp
		435					440					445			
Lys	Ser	Glu	Asn	Leu	His	Phe	Asp	Ala	Phe	Gln	Gly	Phe	Ser	Gln	Gln
	450					455					460				
Tyr	Leu	Val	Glu	Ala	Glu	Val	Phe	Leu	Met	Ser	Ala	Ala	Gln	Val	Gly
465					470					475					480
Leu	Phe	Pro	Lys	Leu	Glu	Leu	Ser	Lys	Arg	Tyr	Pro	Lys	Thr	Phe	Pro
				485					490						495
Phe	Thr	Arg	Ile	Thr	Leu	Asn	Tyr	Phe	Glu	Lys	Arg	Pro	Tyr	Lys	Ile
			500					505					510		
Ser	His	Ala	Tyr	Leu	Ser	Asp	Leu	Pro	Ala	Leu	Val	Asp	Leu	Glu	Val
		515					520					525			
Lys	Cys	Trp	Pro	Glu	Asn	Leu	Arg	Ala	Ser	Thr	His	Glu	Ile	Arg	Arg
	530					535					540				
Arg	Leu	Glu	Leu	Asn	Pro	Gln	Gly	Asn	Leu	Val	Leu	Ile	Ile	Glu	Asp
545					550					555					560
Gln	Ile	Ile	Gly	Ala	Ile	Tyr	Ser	Gln	Thr	Ile	Thr	Ser	Thr	Glu	Ala
				565					570						575
Leu	Glu	Asn	Val	Lys	Tyr	Ala	Gln	Val	Pro	Thr	Leu	His	Thr	Pro	Gln
			580					585					590		
Gly	Ser	Val	Ile	Gln	Leu	Leu	Ala	Leu	Asn	Ile	Leu	Pro	Glu	Phe	Gln
		595					600					605			
Ala	Arg	Gly	Leu	Gly	Asn	Glu	Leu	Arg	Asp	Phe	Met	Leu	Tyr	Tyr	Cys
	610					615					620				
Thr	Leu	Lys	Gly	Gly	Ile	Glu	Ser	Val	Val	Gly	Val	Thr	Arg	Cys	Arg
625					630					635					640
Asn	Tyr	Val	Asn	Tyr	Ser	Gln	Met	Pro	Met	Met	Glu	Tyr	Leu	Lys	Leu
				645					650					655	
His	Asn	Glu	Gln	Arg	Gln	Leu	Leu	Asp	Pro	Ile	Val	Gly	Phe	His	Val
			660					665					670		
Ser	Gly	Gly	Ala	Glu	Ile	Arg	Gly	Ile	Ile	Ala	Asn	Tyr	Arg	Pro	Glu
		675					680					685			
Asp	Thr	Asp	Asn	Leu	Gly	Met	Gly	Ile	Leu	Ile	Glu	Tyr	Asn	Leu	Arg
	690					695					700				
Asp	Ser	Ala	Leu	His	Ser	Pro	Gly	Asp	Arg	Lys	Gly	Pro	Tyr	Ile	Asn
705						710				715					720
Ser	Ala	Ile	Gly	Ser	Leu	Val	Pro	Lys	Ala	Thr	Ser	Ala	Thr	Lys	Glu
				725					730					735	
Asn	Lys	Thr	Val	Ala	Asp	Leu	Val	Lys	Glu	Cys	Ile	Leu	Lys	Val	Met
			740					745					750		
Gly	Ser	Gln	Arg	Gln	Ala	Ala	Tyr	Ala	Pro	Gln	Gln	Lys	Leu	Leu	Asp
		755					760					765			
Met	Gly	Leu	Asp	Ser	Leu	Asp	Leu	Leu	Glu	Leu	Gln	Thr	Leu	Leu	Glu
	770					775					780				
Glu	Arg	Leu	Gly	Ile	Asn	Leu	Ser	Gly	Thr	Phe	Phe	Leu	Gln	Lys	Asn
785					790					795					800
Thr	Pro	Thr	Ala	Ile	Ile	Thr	Tyr	Phe	Gln	Asn	Gln	Val	Val	Gln	Glu
				805					810					815	
Lys	Gln	Ser	Asp	Leu	Ala	Pro	Pro	Val	Asp	Ser	Ala	Asn	Glu	Ile	Asn
			820					825					830		

Thr Leu Glu Asn Val Val Asn Gln Gln Lys Ile Pro Gln Val Thr Arg
 835 840 845
 Val Val Thr Glu Gln Gln Gly Arg Lys Val Leu Ile Asp Gly His Trp
 850 855 860
 Val Ile Asp Phe Ala Ser Cys Asn Tyr Leu Gly Leu Asp Leu His Pro
 865 870 875 880
 Lys Val Lys Glu Ala Ile Pro Pro Ala Leu Asp Lys Trp Gly Thr His
 885 890 895
 Pro Ser Trp Thr Arg Leu Val Ala Ser Pro Ala Ile Tyr Glu Glu Leu
 900 905 910
 Glu Glu Glu Leu Ser Lys Leu Leu Gly Val Pro Asp Val Leu Val Phe
 915 920 925
 Pro Ala Val Thr Leu Leu Gln Ile Gly Ile Leu Pro Leu Leu Thr Gly
 930 935 940
 Asn Asn Gly Val Ile Phe Gly Asp Ile Ala Ala His Arg Cys Ile Tyr
 945 950 955 960
 Glu Ala Cys Cys Leu Ala Gln His Lys Gly Ala Gln Phe Ile Gln Tyr
 965 970 975
 Arg His Asn Asp Leu Asn Asp Leu Ala Glu Lys Leu Ala Lys Tyr Pro
 980 985 990
 Pro Glu Gln Val Lys Ile Ile Val Ile Asp Gly Val Tyr Ser Met Ser
 995 1000 1005
 Ala Asp Phe Pro Asp Leu Pro Ala Tyr Val His Leu Ala Lys Glu
 1010 1015 1020
 Tyr Asn Ala Leu Ile Tyr Met Asp Asp Ala His Gly Phe Gly Ile
 1025 1030 1035
 Leu Gly Glu Asn Pro Ser Ser Asp Met Pro Tyr Gly Tyr Lys Gly
 1040 1045 1050
 Asn Gly Met Val Asn Tyr Phe Asp Leu Arg Phe Ala Glu Asp Asn
 1055 1060 1065
 Ile Ile Tyr Val Ala Gly Leu Ser Lys Ala Tyr Ser Ser Tyr Ala
 1070 1075 1080
 Ala Phe Leu Thr Cys Gly Asp Arg Arg Ile Lys Thr Asn Phe Arg
 1085 1090 1095
 Asn Ala Trp Thr Ala Ile Phe Ser Gly Pro Ser Pro Val Ala Ser
 1100 1105 1110
 Leu Ala Ser Ala Leu Ala Gly Leu Gln Val Asn Arg Gln Glu Gly
 1115 1120 1125
 Glu Gln Leu Arg Lys Gln Ile Tyr His Leu Thr His Lys Leu Val
 1130 1135 1140
 Thr Gln Ala Arg Ala Ile Gly Phe Glu Val Asp Asn Tyr Gly Tyr
 1145 1150 1155
 Val Pro Ile Val Gly Val Leu Val Gly Asp Ala Gln His Met Ile
 1160 1165 1170
 Asp Val Cys Gln Leu Leu Trp Glu Tyr Gly Ile Leu Ile Thr Pro
 1175 1180 1185
 Ala Ile Phe Pro Ile Val Pro Leu Asn Lys Ser Ala Leu Arg Phe
 1190 1195 1200
 Ser Ile Thr Ala Ala Asn Thr Glu Glu Glu Ile Asp Gln Ala Ile
 1205 1210 1215
 Lys Ser Leu Lys Ala Val Trp Asp Leu Leu Gln Lys Arg Lys Ala
 1220 1225 1230
 Leu Pro Cys Lys Gln Glu Glu Asn Ile Leu Lys His
 1235 1240 1245

<210> 16

<211> 387

<212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 16

atgttgaaag atttcaacca gtttttaatc agaacactag cattcgtatt cgcatttggt	60
attttcttaa ccaactggagt tggcattgct aaagctgact acctagttaa aggtggaaag	120
attaccaatg ttcaaaatac ttcttctaac ggtgataatt atgccgttag tatcagcgg	180
gggtttgggtc cttgcgcaga tagagtgatt atcctaccaa cttcaggagt gataaatcga	240
gacattcata tgcgtggcta tgaagccgca ttaactgcac tatccaatgg ctttttagta	300
gatatttacg actatactgg ctcttcttgc agcaatgggtg gccaaactaac tattaccaac	360
caattaggta agctaactcag caattag	387

<210> 17
 <211> 128
 <212> PRT
 <213> *Cylindrospermopsis raciborskii* T3

5

<400> 17

Met	Leu	Lys	Asp	Phe	Asn	Gln	Phe	Leu	Ile	Arg	Thr	Leu	Ala	Phe	Val
1				5					10					15	
Phe	Ala	Phe	Gly	Ile	Phe	Leu	Thr	Thr	Gly	Val	Gly	Ile	Ala	Lys	Ala
			20					25					30		
Asp	Tyr	Leu	Val	Lys	Gly	Gly	Lys	Ile	Thr	Asn	Val	Gln	Asn	Thr	Ser
		35					40					45			
Ser	Asn	Gly	Asp	Asn	Tyr	Ala	Val	Ser	Ile	Ser	Gly	Gly	Phe	Gly	Pro
	50				55					60					
Cys	Ala	Asp	Arg	Val	Ile	Ile	Leu	Pro	Thr	Ser	Gly	Val	Ile	Asn	Arg
65					70					75				80	
Asp	Ile	His	Met	Arg	Gly	Tyr	Glu	Ala	Ala	Leu	Thr	Ala	Leu	Ser	Asn
				85					90					95	
Gly	Phe	Leu	Val	Asp	Ile	Tyr	Asp	Tyr	Thr	Gly	Ser	Ser	Cys	Ser	Asn
			100					105					110		
Gly	Gly	Gln	Leu	Thr	Ile	Thr	Asn	Gln	Leu	Gly	Lys	Leu	Ile	Ser	Asn
		115					120					125			

10

<210> 18
 <211> 1416
 <212> ADN
 <213> *Cylindrospermopsis raciborskii* T3

15

<400> 18

atggaacaa	cctcaaaaa	atttaagtca	gatctgatat	tagaagcacg	agcaagccta	60
aagttgggaa	tccccttagt	catttcacaa	atgtgcgaaa	cgggtattta	tacagcgaat	120
gcagtcata	tgggtttact	tggtacgcaa	gttttggccg	ccggtgcttt	gggcgcgctc	180
gcttttttga	ccttattatt	tgctgccat	ggatttctct	cagtaggagg	atcactagca	240
gccgaagctt	ttggggcaaa	taaaatagat	gaagttagtc	gtattgcttc	cgggcaata	300
tggctagcag	ttacctgtgc	tttacctgca	atgcttctgc	tttggcatgg	cgatactatc	360
ttgctgctat	tcgggtcaaga	ggaaagcaat	gtgttattga	caaaaacgta	tttactactca	420
attttatggg	gctttccgc	tgcgcttagt	attttgacat	taagaggcat	tgctctgct	480
ctcaacgttc	cccgaattgat	aactattact	atgctcactc	agctgatatt	gaataccgcc	540
gccgattatg	tgtaaatatt	cggtaaattt	ggtcttcctc	aacttggttt	ggctggaata	600
ggctgggcaa	ctgctctggg	tttttgggtt	agttttacat	tggggcttat	cttgctgatt	660
ttctccctga	aagttagaga	ttataaactt	ttccgctact	tgcatcagtt	tgataaacag	720
atctttgtca	aaatttttca	aactggatgg	cccatggggt	ttcaatgggg	ggcggaaacg	780
gcactattta	acgtcaccgc	ttgggtagca	gggtatttag	gaacggtaac	attagcagcc	840
catgatattg	gcttccaaac	ggcagaactg	gcgatggtta	taccactcgg	agtcggcaat	900
gtcgctatga	caagagtagg	tcagagtata	ggagaaaaaa	accctttggg	tgcaagaagg	960
gtagcatcga	ttggaattac	aatagttggc	atttatgcc	gtattgtagc	acttggtttc	1020
tggttgtttc	catatcaaat	tgccggaatt	tatttaaata	taaacaatcc	cgagaatatc	1080
gaagcaatta	agaaagcaac	tacttttatc	cccttggcgg	gactattcca	aatgttttac	1140
agttattcaa	taattattgt	tggggctttg	gtcgggtctgc	gggatacatt	tgttccagta	1200
tcaatgaact	taattgtctg	gggtcttgga	ttggcaggaa	gctatttcat	ggcaatcatt	1260
ttaggatggg	gggggatcgg	gatttgggtg	gctatgggtt	tgagtccact	cctctcggca	1320
gttattttta	ctgttcgttt	ttatcgagtg	attgacaatc	ttcttgccaa	cagtgatgat	1380
atgttacaga	atgcgtctgt	tactactcta	ggctga			1416

20

<210> 19
 <211> 471
 <212> PRT
 <213> *Cylindrospermopsis raciborskii* T3

25

<400> 19

```

Met Glu Thr Thr Ser Lys Lys Phe Lys Ser Asp Leu Ile Leu Glu Ala
1      5      10      15
Arg Ala Ser Leu Lys Leu Gly Ile Pro Leu Val Ile Ser Gln Met Cys
      20      25      30
Glu Thr Gly Ile Tyr Thr Ala Asn Ala Val Met Met Gly Leu Leu Gly
      35      40      45
Thr Gln Val Leu Ala Ala Gly Ala Leu Gly Ala Leu Ala Phe Leu Thr
      50      55      60
Leu Leu Phe Ala Cys His Gly Ile Leu Ser Val Gly Gly Ser Leu Ala
65      70      75      80
Ala Glu Ala Phe Gly Ala Asn Lys Ile Asp Glu Val Ser Arg Ile Ala
      85      90      95
Ser Gly Gln Ile Trp Leu Ala Val Thr Leu Ser Leu Pro Ala Met Leu
      100     105     110
Leu Leu Trp His Gly Asp Thr Ile Leu Leu Leu Phe Gly Gln Glu Glu
      115     120     125
Ser Asn Val Leu Leu Thr Lys Thr Tyr Leu His Ser Ile Leu Trp Gly
      130     135     140
Phe Pro Ala Ala Leu Ser Ile Leu Thr Leu Arg Gly Ile Ala Ser Ala
145     150     155     160
Leu Asn Val Pro Arg Leu Ile Thr Ile Thr Met Leu Thr Gln Leu Ile
      165     170     175
Leu Asn Thr Ala Ala Asp Tyr Val Leu Ile Phe Gly Lys Phe Gly Leu
      180     185     190
Pro Gln Leu Gly Leu Ala Gly Ile Gly Trp Ala Thr Ala Leu Gly Phe
      195     200     205
Trp Val Ser Phe Thr Leu Gly Leu Ile Leu Leu Ile Phe Ser Leu Lys
      210     215     220
Val Arg Asp Tyr Lys Leu Phe Arg Tyr Leu His Gln Phe Asp Lys Gln
225     230     235     240
Ile Phe Val Lys Ile Phe Gln Thr Gly Trp Pro Met Gly Phe Gln Trp
      245     250     255
Gly Ala Glu Thr Ala Leu Phe Asn Val Thr Ala Trp Val Ala Gly Tyr
      260     265     270
Leu Gly Thr Val Thr Leu Ala Ala His Asp Ile Gly Phe Gln Thr Ala
      275     280     285
Glu Leu Ala Met Val Ile Pro Leu Gly Val Gly Asn Val Ala Met Thr
      290     295     300
Arg Val Gly Gln Ser Ile Gly Glu Lys Asn Pro Leu Gly Ala Arg Arg
305     310     315     320
Val Ala Ser Ile Gly Ile Thr Ile Val Gly Ile Tyr Ala Ser Ile Val
      325     330     335
Ala Leu Val Phe Trp Leu Phe Pro Tyr Gln Ile Ala Gly Ile Tyr Leu
      340     345     350
Asn Ile Asn Asn Pro Glu Asn Ile Glu Ala Ile Lys Lys Ala Thr Thr
355     360     365

Phe Ile Pro Leu Ala Gly Leu Phe Gln Met Phe Tyr Ser Ile Gln Ile
370     375     380
Ile Ile Val Gly Ala Leu Val Gly Leu Arg Asp Thr Phe Val Pro Val
385     390     395     400
Ser Met Asn Leu Ile Val Trp Gly Leu Gly Leu Ala Gly Ser Tyr Phe
      405     410     415
Met Ala Ile Ile Leu Gly Trp Gly Gly Ile Gly Ile Trp Leu Ala Met
      420     425     430
Val Leu Ser Pro Leu Leu Ser Ala Val Ile Leu Thr Val Arg Phe Tyr
      435     440     445
Arg Val Ile Asp Asn Leu Leu Ala Asn Ser Asp Asp Met Leu Gln Asn
      450     455     460
Ala Ser Val Thr Thr Leu Gly
465     470

```

5 <210> 20
 <211> 1134
 <212> ADN
 <213> Cylindrospermopsis raciborskii T3

10 <400> 20

```

atgaccaatc aaaataacca agaattagag aacgatttac caatcgccaa gcagccttgt      60
ccgggtcaatt cttataatga gtgggacaca cttgaggagg tcattgttgg tagtggtgaa      120
ggtgcaatgt taccggccct agaaccaatc aacaaatgga cattcccttt tgaagaattg      180
gaatctgccc aaaagatact ctctgagagg ggaggagttc cttatccacc agagatgatt      240
acattagcac acaaagaact aaatgaattt attcacattc ttgaagcaga aggggtcaaa      300
gttcgtcgag ttaaacctgt agatttctct gtccccttct ccacaccagc ttggcaagta      360
ggaagtgggt tttgtgccgc caatcctcgc gatgtttttt tgggtattgg gaatgagatt      420
attgaagcac caatggcaga tcgcaaccgc tattttgaaa cttgggcgta tcgagagatg      480
ctcaaggaat attttcaggc aggagctaag tggactgcag cgccgaagcc acaattattc      540
gacgcacagt atgacttcaa tttccagttt cttcaactgg gggagccgcc gcgtttcgtc      600
gttacagagt ttgaaccgac ttttgatgcg gcagattttg tgcgctgtgg acgagatatt      660
tttggtcaaa aaagtcatgt gactaatggt ttgggcatag aatggttaca acgtcacttg      720
gaagacgaat accgtattca tattattgaa tcgcattgtc cggaagcact gcacatcgat      780
accaccttaa tgcctcttgc acctggcaaa atactagtaa atccagaatt tgtagatggt      840
aataaattgc caaaaatcct gaaaagctgg gacattttgg ttgcacctta cccaaccat      900
atactcaaaa accagctgag actggtcagt gaatgggcag gtttgaatgt actgatgtta      960
gatgaagagc gagtcattgt agaaaaaac caggagcaga tgattaaagc actgaaagat     1020
tggggattta agcctattgt ttgccatttt gaaagctact atccattttt aggatcattt     1080
cactgtgcaa cattagacgt tcgccgacgc ggaactcttc agtcctattt ttaa      1134

```

<210> 21

<211> 377

5 <212> PRT

<213> Cylindrospermopsis raciborskii T3

<400> 21

```

Met Thr Asn Gln Asn Asn Gln Glu Leu Glu Asn Asp Leu Pro Ile Ala
1                               5                               10                               15
Lys Gln Pro Cys Pro Val Asn Ser Tyr Asn Glu Trp Asp Thr Leu Glu
20                               25                               30
Glu Val Ile Val Gly Ser Val Glu Gly Ala Met Leu Pro Ala Leu Glu
35                               40                               45
Pro Ile Asn Lys Trp Thr Phe Pro Phe Glu Glu Leu Glu Ser Ala Gln
50                               55                               60
Lys Ile Leu Ser Glu Arg Gly Gly Val Pro Tyr Pro Pro Glu Met Ile
65                               70                               75                               80
Thr Leu Ala His Lys Glu Leu Asn Glu Phe Ile His Ile Leu Glu Ala
85                               90                               95

```

10

Glu Gly Val Lys Val Arg Arg Val Lys Pro Val Asp Phe Ser Val Pro
100 105 110
Phe Ser Thr Pro Ala Trp Gln Val Gly Ser Gly Phe Cys Ala Ala Asn
115 120 125
Pro Arg Asp Val Phe Leu Val Ile Gly Asn Glu Ile Ile Glu Ala Pro
130 135 140
Met Ala Asp Arg Asn Arg Tyr Phe Glu Thr Trp Ala Tyr Arg Glu Met
145 150 155 160
Leu Lys Glu Tyr Phe Gln Ala Gly Ala Lys Trp Thr Ala Ala Pro Lys
165 170 175
Pro Gln Leu Phe Asp Ala Gln Tyr Asp Phe Asn Phe Gln Phe Pro Gln
180 185 190
Leu Gly Glu Pro Pro Arg Phe Val Val Thr Glu Phe Glu Pro Thr Phe
195 200 205
Asp Ala Ala Asp Phe Val Arg Cys Gly Arg Asp Ile Phe Gly Gln Lys
210 215 220
Ser His Val Thr Asn Gly Leu Gly Ile Glu Trp Leu Gln Arg His Leu
225 230 235 240
Glu Asp Glu Tyr Arg Ile His Ile Ile Glu Ser His Cys Pro Glu Ala
245 250 255
Leu His Ile Asp Thr Thr Leu Met Pro Leu Ala Pro Gly Lys Ile Leu
260 265 270
Val Asn Pro Glu Phe Val Asp Val Asn Lys Leu Pro Lys Ile Leu Lys
275 280 285
Ser Trp Asp Ile Leu Val Ala Pro Tyr Pro Asn His Ile Pro Gln Asn
290 295 300
Gln Leu Arg Leu Val Ser Glu Trp Ala Gly Leu Asn Val Leu Met Leu
305 310 315 320
Asp Glu Glu Arg Val Ile Val Glu Lys Asn Gln Glu Gln Met Ile Lys
325 330 335
Ala Leu Lys Asp Trp Gly Phe Lys Pro Ile Val Cys His Phe Glu Ser
340 345 350
Tyr Tyr Pro Phe Leu Gly Ser Phe His Cys Ala Thr Leu Asp Val Arg
355 360 365
Arg Arg Gly Thr Leu Gln Ser Tyr Phe
370 375

<210> 22

<211> 1005

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 22

atgacaactg ctgacctaat ctttaattaac aactgggtacg tagtcgcaaa ggtggaagat 60
tgtaaaccag gaagtatcac cacggctctt ttattggggag ttaagttggt actatggcgc 120
agtcgtgaac agaattcccc catacagata tggcaagact actgccctca ccgaggtgtg 180
gctctgtceta tgggagaaat tgtaataat actttgggtt gtccgtatca cggatggaga 240
tataatcaag caggtaaatt cgtacatatc ccggctcacc ctgacatgac acccccagca 300
agtgcccaag ccaagatcta tcattgccag gacgatacg gattagtatg ggtgtgctta 360
gggtatcctg tcaatgatat accttcatta cccgaatggg acgatccgaa ttatcataat 420
acttgtacta aatcttattt tattcaagct agtgcgtttc gtgtaattgga taatttcata 480
gatgtatctc attttccttt tgtccacgac ggtgggttag gtgatcgcaa ccacgcacaa 540
attgaagaat ttgaggtaaa agtagacaaa gatggcatta gcataggtaa ccttaaaactc 600
cagatgccaa ggtttaacag cagtaacgaa gatgactcat ggactcttta ccaaaggatt 660
agtcacccct tgtgtcaata ctatattact gaatcctctg aaattcggac tgcggatttg 720
atgctggtaa caccgattga tgaagacaac agcttagtgc gaatgttagt aacgtggaac 780
cgctccgaaa tattagagtc aacggtacta gaggaatttg acgaaacaat agaacaagat 840
attccgatta tacactctca acagccagcg cgtttaccac tgttaccttc aaagcagata 900

aacatgcaat ggttgtcaca ggaaatacat gtaccgtcag atcgatgcac agttgcctat 960
cgtcgatggc taaaggaact gggcgttacc tatgggtgtt gtttaa 1005

<210> 23

15 <211> 334

<212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 23

20

Met Thr Thr Ala Asp Leu Ile Leu Ile Asn Asn Trp Tyr Val Val Ala
1 5 10 15
Lys Val Glu Asp Cys Lys Pro Gly Ser Ile Thr Thr Ala Leu Leu Leu
20 25 30
Gly Val Lys Leu Val Leu Trp Arg Ser Arg Glu Gln Asn Ser Pro Ile
35 40 45
Gln Ile Trp Gln Asp Tyr Cys Pro His Arg Gly Val Ala Leu Ser Met
50 55 60
Gly Glu Ile Val Asn Asn Thr Leu Val Cys Pro Tyr His Gly Trp Arg
65 70 75 80
Tyr Asn Gln Ala Gly Lys Cys Val His Ile Pro Ala His Pro Asp Met
85 90 95
Thr Pro Pro Ala Ser Ala Gln Ala Lys Ile Tyr His Cys Gln Glu Arg
100 105 110
Tyr Gly Leu Val Trp Val Cys Leu Gly Asp Pro Val Asn Asp Ile Pro
115 120 125
Ser Leu Pro Glu Trp Asp Asp Pro Asn Tyr His Asn Thr Cys Thr Lys
130 135 140
Ser Tyr Phe Ile Gln Ala Ser Ala Phe Arg Val Met Asp Asn Phe Ile
145 150 155 160
Asp Val Ser His Phe Pro Phe Val His Asp Gly Gly Leu Gly Asp Arg
165 170 175
Asn His Ala Gln Ile Glu Glu Phe Glu Val Lys Val Asp Lys Asp Gly
180 185 190
Ile Ser Ile Gly Asn Leu Lys Leu Gln Met Pro Arg Phe Asn Ser Ser
195 200 205
Asn Glu Asp Asp Ser Trp Thr Leu Tyr Gln Arg Ile Ser His Pro Leu
210 215 220
Cys Gln Tyr Tyr Ile Thr Glu Ser Ser Glu Ile Arg Thr Ala Asp Leu
225 230 235 240
Met Leu Val Thr Pro Ile Asp Glu Asp Asn Ser Leu Val Arg Met Leu
245 250 255
Val Thr Trp Asn Arg Ser Glu Ile Leu Glu Ser Thr Val Leu Glu Glu
260 265 270
Phe Asp Glu Thr Ile Glu Gln Asp Ile Pro Ile Ile His Ser Gln Gln
275 280 285
Pro Ala Arg Leu Pro Leu Leu Pro Ser Lys Gln Ile Asn Met Gln Trp
290 295 300
Leu Ser Gln Glu Ile His Val Pro Ser Asp Arg Cys Thr Val Ala Tyr
305 310 315 320
Arg Arg Trp Leu Lys Glu Leu Gly Val Thr Tyr Gly Val Cys
325 330

<210> 24

<211> 1839

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 24

10 atgcagatct taggaatttc agcttactac cagcatagtg ctgccgcgat gggtatcgat 60

```

ggcgaaattg ttgctgcagc tcaggaagaa cgtttctcaa gacgaaagca cgatgctggg 120
tttccgactg gagcgattac ttactgtcta aaacaagtag gaaccaagtt acaatatatc 180
gatcaaatg ttttttacga caagccatta gtcaaatttg agcgggttgct agaaacatat 240
ttagcatatg ccccaaaggg atttggctcg tttattactg ctatgcccgt ttgggtcaaa 300
gaaaagcttt acctaaaaac acttttaaaa aaagaattgg cgcttttggg ggagtgc aaa 360
gcttttcaat tgctctctct actgtttacc tcacatcacc aagcccatgc ggccgctgct 420
ttttttccca gtcccttttca gcgtgctgcc gttctgtgct tagatgggtg aggagagtgg 480
gcaactactt ctgtctgggt gggagaagga aataaactca caccacaatg ggaaattgat 540
tttccccatt ccctcgggtt gctttactca gcgtttacct actacactgg gttcaaagtt 600
aactcagggt agtacaaact catgggttta gcacctacg gggaacccaa atatgtggac 660
caaatttctca agcatttgtt ggatctcaaa gaagatggta cttttagggtt gaatatggac 720
tacttcaact acacgggtgg gctaaccatg accaatcata agttccatag tatgtttgga 780
ggaccaccac gccaggcgga aggaaaaatc tcccaaagag acatggatct ggcaagttcg 840
atccaaaagg tgactgaaga agtcatactg cgtctggcta gaactatcaa aaaagaactg 900
gggtgtagagt atctatgttt agcaggtggt gtccgtctca attgcgtggc taacggacga 960
attctccgag aaagtgattt caaagatatt tggattcaac ccgcagcagg agatgccggt 1020
agtgcagtgg gagcagcttt agcgatttgg catgaatacc ataagaaacc tcgcacttca 1080
acagcaggcg atcgcgatga aggttcttat ctgggaccta gctttagcga ggccggagatt 1140
ctccagtttc ttaattctgt taacataccc taccatcgat gcgttgataa cgaacttatg 1200
gctcgtcttt cagaaatttt agaccagga aatgtttag gctgggtttc tggacgaatg 1260
gagtttggtc cgcgtgcttt ggggtggcgt tcgattattg gcgattcacg cagtccaaaa 1320
atgcaatcgg tcatgaacct gaaaattaaa tatcgtgagt ccttcggtcc atttgctcct 1380
tcagtcttgg ctgaacgagt ctccgactac ttcgatcttg atcgtcctag tccttatatg 1440
cttttggtag cacaagtcaa agagaatctg cacattccta tgacacaaga gcaacacgag 1500
ctatttggga tcgagaagct gaatttctct cgttcccaaa ttcccgagc cactcacgtt 1560
gattactcag ctcgatttca gacagttcac aaagaaacga atcctcgta ctacgagtta 1620
attcgtcatt ttgaggcacg aactggttgt gctgtcttgg tcaatacttc gtttaattgtc 1680
cgcggcgaac caattgtttg tactcccgaa gacgcttatc gatgctttat gagaactgaa 1740
atggactatt tgggtatgga gaatttcttg ttggtcaaat ctgaacagcc acggggaaat 1800
agtgatgagt catggcaaaa agaattcgag ttagattaa 1839

```

<210> 25

<211> 612

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 25

```

Met Gln Ile Leu Gly Ile Ser Ala Tyr Tyr His Asp Ser Ala Ala Ala
1      5      10      15
Met Val Ile Asp Gly Glu Ile Val Ala Ala Ala Gln Glu Glu Arg Phe
20     25     30
Ser Arg Arg Lys His Asp Ala Gly Phe Pro Thr Gly Ala Ile Thr Tyr
35     40     45
Cys Leu Lys Gln Val Gly Thr Lys Leu Gln Tyr Ile Asp Gln Ile Val
50     55     60
Phe Tyr Asp Lys Pro Leu Val Lys Phe Glu Arg Leu Leu Glu Thr Tyr
65     70     75     80
Leu Ala Tyr Ala Pro Lys Gly Phe Gly Ser Phe Ile Thr Ala Met Pro
85     90     95
Val Trp Leu Lys Glu Lys Leu Tyr Leu Lys Thr Leu Leu Lys Lys Glu
100    105    110
Leu Ala Leu Leu Gly Glu Cys Lys Ala Ser Gln Leu Pro Pro Leu Leu
115    120    125
Phe Thr Ser His His Gln Ala His Ala Ala Ala Ala Phe Phe Pro Ser
130    135    140
Pro Phe Gln Arg Ala Ala Val Leu Cys Leu Asp Gly Val Gly Glu Trp
145    150    155    160
Ala Thr Thr Ser Val Trp Leu Gly Glu Gly Asn Lys Leu Thr Pro Gln

```

10

				165					170					175	
Trp	Glu	Ile	Asp	Phe	Pro	His	Ser	Leu	Gly	Leu	Leu	Tyr	Ser	Ala	Phe
			180					185					190		
Thr	Tyr	Tyr	Thr	Gly	Phe	Lys	Val	Asn	Ser	Gly	Glu	Tyr	Lys	Leu	Met
			195				200					205			
Gly	Leu	Ala	Pro	Tyr	Gly	Glu	Pro	Lys	Tyr	Val	Asp	Gln	Ile	Leu	Lys
			210			215					220				
His	Leu	Leu	Asp	Leu	Lys	Glu	Asp	Gly	Thr	Phe	Arg	Leu	Asn	Met	Asp
225				230						235					240
Tyr	Phe	Asn	Tyr	Thr	Val	Gly	Leu	Thr	Met	Thr	Asn	His	Lys	Phe	His
				245					250					255	
Ser	Met	Phe	Gly	Gly	Pro	Pro	Arg	Gln	Ala	Glu	Gly	Lys	Ile	Ser	Gln
			260					265					270		
Arg	Asp	Met	Asp	Leu	Ala	Ser	Ser	Ile	Gln	Lys	Val	Thr	Glu	Glu	Val
			275				280						285		
Ile	Leu	Arg	Leu	Ala	Arg	Thr	Ile	Lys	Lys	Glu	Leu	Gly	Val	Glu	Tyr
			290			295					300				
Leu	Cys	Leu	Ala	Gly	Gly	Val	Gly	Leu	Asn	Cys	Val	Ala	Asn	Gly	Arg
305				310						315					320
Ile	Leu	Arg	Glu	Ser	Asp	Phe	Lys	Asp	Ile	Trp	Ile	Gln	Pro	Ala	Ala
				325					330					335	
Gly	Asp	Ala	Gly	Ser	Ala	Val	Gly	Ala	Ala	Leu	Ala	Ile	Trp	His	Glu
			340					345					350		
Tyr	His	Lys	Lys	Pro	Arg	Thr	Ser	Thr	Ala	Gly	Asp	Arg	Met	Lys	Gly
			355				360					365			
Ser	Tyr	Leu	Gly	Pro	Ser	Phe	Ser	Glu	Ala	Glu	Ile	Leu	Gln	Phe	Leu
			370			375					380				
Asn	Ser	Val	Asn	Ile	Pro	Tyr	His	Arg	Cys	Val	Asp	Asn	Glu	Leu	Met
385				390						395					400
Ala	Arg	Leu	Ala	Glu	Ile	Leu	Asp	Gln	Gly	Asn	Val	Val	Gly	Trp	Phe
				405					410					415	
Ser	Gly	Arg	Met	Glu	Phe	Gly	Pro	Arg	Ala	Leu	Gly	Gly	Arg	Ser	Ile
			420					425					430		
Ile	Gly	Asp	Ser	Arg	Ser	Pro	Lys	Met	Gln	Ser	Val	Met	Asn	Leu	Lys
			435				440					445			
Ile	Lys	Tyr	Arg	Glu	Ser	Phe	Arg	Pro	Phe	Ala	Pro	Ser	Val	Leu	Ala
					455					460					
Glu	Arg	Val	Ser	Asp	Tyr	Phe	Asp	Leu	Asp	Arg	Pro	Ser	Pro	Tyr	Met
465				470						475					480
Leu	Leu	Val	Ala	Gln	Val	Lys	Glu	Asn	Leu	His	Ile	Pro	Met	Thr	Gln
				485					490					495	
Glu	Gln	His	Glu	Leu	Phe	Gly	Ile	Glu	Lys	Leu	Asn	Val	Pro	Arg	Ser
			500					505					510		
Gln	Ile	Pro	Ala	Val	Thr	His	Val	Asp	Tyr						

610
 <210> 26
 <211> 444
 <212> ADN
 <213> *Cylindrospermopsis raciborskii* T3

<400> 26

```

atgagtgaaat ttttcccaca aaaaagtgggt aaattaaaga tggaacagat aaaagaactt      60
gacaaaaaag gattgctga gtttggactg attggcgggt ctatagtggc ggttttattc      120
ggctttttac tgccagttat acgccatcat tccttatcag ttatcccttg ggttgttgct      180
ggattttctt ggatttgggc aataatcgca cctacgactt taagttttat ttaccaaata      240
tggatgagga ttggacttgt tttaggatgg atacaaacac gaattatttt gggagtttta      300
ttttatataa tgatcacacc aataggattc ataagacggc tgttgaatca agatccaatg      360
acgcgaatct tcgagccaga gttgccaact tatcgccaat tgagtaagtc aagaactaca      420
caaagtatgg agaaccatt ctaa                                     444

```

<210> 27

<211> 147

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 27

```

Met Ser Glu Phe Phe Pro Gln Lys Ser Gly Lys Leu Lys Met Glu Gln
1      5      10      15
Ile Lys Glu Leu Asp Lys Lys Gly Leu Arg Glu Phe Gly Leu Ile Gly
      20      25      30

Gly Ser Ile Val Ala Val Leu Phe Gly Phe Leu Leu Pro Val Ile Arg
      35      40      45
His His Ser Leu Ser Val Ile Pro Trp Val Val Ala Gly Phe Leu Trp
      50      55      60
Ile Trp Ala Ile Ile Ala Pro Thr Thr Leu Ser Phe Ile Tyr Gln Ile
65      70      75      80
Trp Met Arg Ile Gly Leu Val Leu Gly Trp Ile Gln Thr Arg Ile Ile
      85      90      95
Leu Gly Val Leu Phe Tyr Ile Met Ile Thr Pro Ile Gly Phe Ile Arg
      100      105      110
Arg Leu Leu Asn Gln Asp Pro Met Thr Arg Ile Phe Glu Pro Glu Leu
      115      120      125
Pro Thr Tyr Arg Gln Leu Ser Lys Ser Arg Thr Thr Gln Ser Met Glu
      130      135      140
Lys Pro Phe
145

```

10

<210> 28

<211> 165

<212> ADN

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 28

```

atgctaaaag acacttggga ttttattaaa gacattgccg gatttattaa agaacaaaaa      60
aactatttgt tgattcccct aattatcacc ctggtatcct tgggggcgct gattgtcttt      120
gctcaatctt ctgcgatcgc acctttcatt tacactcttt ttttaa                                     165

```

20

<210> 29

<211> 54

<212> PRT

<213> *Cylindrospermopsis raciborskii* T3

25

<400> 29

```

Met Leu Lys Asp Thr Trp Asp Phe Ile Lys Asp Ile Ala Gly Phe Ile
1      5      10      15
Lys Glu Gln Lys Asn Tyr Leu Leu Ile Pro Leu Ile Ile Thr Leu Val
      20      25      30
Ser Leu Gly Ala Leu Ile Val Phe Ala Gln Ser Ser Ala Ile Ala Pro
      35      40      45
Phe Ile Tyr Thr Leu Phe
      50

```

30

<210> 30

<211> 1299

<212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 30

```

atgagtaact tcaaggggttc ggtaaagata gcattgatgg gaatattgat tttttgtggg 60
ctaattcttg gcgtagcatt tggtgaaatt gggttacgta ttgccgggat cgaacacata 120
gcattccata gcattgatga acacaggggg tgggtagggc gacctcatgt ttccgggtgg 180
tatagaaccg aaggtgaagc tcacatccaa atgaatagt atggctttcg agatcgagaa 240
cacatcaagg tcaaaccaga aaataccttc aggatagcgc tgttgggaga ttccctttgta 300
gagtccatgc aagtaccgtt ggagcaaaat ttggcagcag ttatagaagg agaaatcagt 360
agttgtatag ctttagctgg acgaaaggcg gaagtgatta attttggagt gactgggttat 420
ggaacagacc aagaactaat tactctacgg gagaaagtgt gggactattc acctgatata 480
gtagtgc tagtatttatac tggcaacgac attgttgata actcccgtgc gctgagtcag 540
aaattctatc ctaatgaact aggttcacta aagccgtttt ttatacttag agatggtaat 600
ctggtgggtt atgcttctgt tatcaatacg gataattatc gctcaaagct gacatgggtg 660
ggcaaaactt atatgaaaat aaaagaccac tcacggattt tacaggtttt aaacatggta 720
cgggatgctc ttaacaactc tagtagaggg ttttcttctc aagctataga ggaaccgtta 780
tttagtgatg gaaaacagga taaaaaattg agcgggtttt ttgatattcta caaaccacct 840
actgaccctg aatggcaaca ggcatggcaa gtcacagaga aactgattag ctcaatgcaa 900
cacgagggtg ctgcgaagaa agcagatttt ttagttgtta cttttggcgg tccctttcaa 960
cgagaacctt tagtgctgca aaaagaaatg caagaattgg gtctgactga ttggttttac 1020
ccagagaagc gaattacacg tttgggtgag gatgaggggt tcagtgtact caatctcagc 1080
ccaaatttgc aggtttattc tgagcagaac aatgcttgcc tatatgggtt tgatgatact 1140
caaggctgtg tagggcattg gaatgcttta ggacatcagg tagcaggaaa aatgattgca 1200
tcgaagattt gtcaacagca gatgagagaa agtatattgc ctcataagca cgacccttca 1260
agccaaagct cacctattac ccaatcagtg atccaataa 1299

```

5

<210> 31

<211> 432

<212> PRT

<213> *Cylindrospermopsis raciborskii* T3

10

<400> 31

```

Met Ser Asn Phe Lys Gly Ser Val Lys Ile Ala Leu Met Gly Ile Leu
1          5          10          15
Ile Phe Cys Gly Leu Ile Phe Gly Val Ala Phe Val Glu Ile Gly Leu
20          25          30
Arg Ile Ala Gly Ile Glu His Ile Ala Phe His Ser Ile Asp Glu His
35          40          45
Arg Gly Trp Val Gly Arg Pro His Val Ser Gly Trp Tyr Arg Thr Glu
50          55          60
Gly Glu Ala His Ile Gln Met Asn Ser Asp Gly Phe Arg Asp Arg Glu
65          70          75          80
His Ile Lys Val Lys Pro Glu Asn Thr Phe Arg Ile Ala Leu Leu Gly
85          90          95
Asp Ser Phe Val Glu Ser Met Gln Val Pro Leu Glu Gln Asn Leu Ala
100         105         110
Ala Val Ile Glu Gly Glu Ile Ser Ser Cys Ile Ala Leu Ala Gly Arg

```

115	120	125
Lys Ala Glu Val Ile Asn Phe Gly Val Thr Gly Tyr Gly Thr Asp Gln		
130	135	140
Glu Leu Ile Thr Leu Arg Glu Lys Val Trp Asp Tyr Ser Pro Asp Ile		
145	150	155
Val Val Leu Asp Phe Tyr Thr Gly Asn Asp Ile Val Asp Asn Ser Arg		
165	170	175
Ala Leu Ser Gln Lys Phe Tyr Pro Asn Glu Leu Gly Ser Leu Lys Pro		
180	185	190
Phe Phe Ile Leu Arg Asp Gly Asn Leu Val Val Asp Ala Ser Phe Ile		
195	200	205
Asn Thr Asp Asn Tyr Arg Ser Lys Leu Thr Trp Trp Gly Lys Thr Tyr		
210	215	220
Met Lys Ile Lys Asp His Ser Arg Ile Leu Gln Val Leu Asn Met Val		
225	230	235
Arg Asp Ala Leu Asn Ser Ser Arg Gly Phe Ser Ser Gln Ala Ile		
245	250	255
Glu Glu Pro Leu Phe Ser Asp Gly Lys Gln Asp Thr Lys Leu Ser Gly		
260	265	270
Phe Phe Asp Ile Tyr Lys Pro Pro Thr Asp Pro Glu Trp Gln Gln Ala		
275	280	285
Trp Gln Val Thr Glu Lys Leu Ile Ser Ser Met Gln His Glu Val Thr		
290	295	300
Ala Lys Lys Ala Asp Phe Leu Val Val Thr Phe Gly Gly Pro Phe Gln		
305	310	315
Arg Glu Pro Leu Val Arg Gln Lys Glu Met Gln Glu Leu Gly Leu Thr		
325	330	335
Asp Trp Phe Tyr Pro Glu Lys Arg Ile Thr Arg Leu Gly Glu Asp Glu		
340	345	350
Gly Phe Ser Val Leu Asn Leu Ser Pro Asn Leu Gln Val Tyr Ser Glu		
355	360	365
Gln Asn Asn Ala Cys Leu Tyr Gly Phe Asp Asp Thr Gln Gly Cys Val		
370	375	380
Gly His Trp Asn Ala Leu Gly His Gln Val Ala Gly Lys Met Ile Ala		
385	390	395
Ser Lys Ile Cys Gln Gln Met Arg Glu Ser Ile Leu Pro His Lys		
405	410	415
His Asp Pro Ser Ser Gln Ser Ser Pro Ile Thr Gln Ser Val Ile Gln		
420	425	430

<210> 32

<211> 1449

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 32

atgacaaata	ccgaaagagg	attagcagaa	ataacatcaa	caggatataa	gtcagagctt	60
agatcggagg	cacgagttag	cctccaactg	gcaattccct	tagtccttgt	cgaaatatgc	120
ggaacgagta	ttaatgtggt	ggatgtagtc	atgatgggct	tacttggtac	tcaagttttg	180
gctgctggtg	ccttggggtgc	gacgcttttt	ttatctgtat	cgaatacttg	ttataatatg	240
cttttgtcgg	gggtagcaaa	ggcatctgag	gcttttgggg	caaacaaaat	agatcaggtt	300
agtcgtattg	cttctgggca	aatatggctg	gcactcacct	tgtctttgcc	tgcaatgctt	360
ttgctttggt	atatggatac	tatattggtg	ctatttggtc	aagttgaaag	caacacatta	420
attgcaaaaa	cgtattttaca	ctcaattgtg	tggggatttc	cggcggcagt	tggtattttg	480
atattaagag	gcattgcctc	tgctgtgaac	gtcccccaat	tggttaactgt	gacgatgcta	540
gtagggctgg	tcttgaatgc	cccggccaat	tatgtattaa	tggtcggtaa	atttggtctt	600
cctgaacttg	gttttagctgg	aataggctgg	gcaagtactt	tggttttttg	gattagtttt	660
ctagtggggg	ttgtcttgct	gattttctcc	ccaaaagtta	gagattataa	acttttccgc	720

10

```

tacttgcatac agtttgatcg acagacgggtt gtggaaattht ttcaaactgg atggcctatg 780
ggtttttctac tgggagtggga atcagtagta ttgagcctca ccgcttggtt aacaggctat 840
ttgggaacag taacattagc agctcatgag atcgcgatcc aaacagcaga actggcgata 900
gtgataccac tcggaatcgg gaatgttgcc gtcacgagag taggtcagac tataggagaa 960
aaaaaccctt tgggtgctag aagggcagca ttgattggga ttatgattgg tggcatttat 1020
gccagtcttg tggcagtcac tttctgggtg tttccatata agattgcggg actttattta 1080
aaaataaacg atccagagag tatggaagca gttgaagacag caactaattt tctcttcttg 1140
gcgggattat tccaattttt tcatagcgtt caaataattg ttgttggggg tttaataggg 1200
ttgcaggata cgtttatccc attgttaatg aatttggtag gctgggggtct tggcttgga 1260
gtaagctatt acatgggaat cattttatgt tggggaggta tgggtatctg gttaggtctg 1320
gttttgagtc cactcctgtc cggacttatt ttaatgggtc gtttttatca agagattgcc 1380
aataggattg ccaatagtga tgatgggcaa gagagtatat ctattgacaa cgttgaagaa 1440
ctctcctga 1449

```

<210> 33

<211> 482

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 33

```

Met Thr Asn Thr Glu Arg Gly Leu Ala Glu Ile Thr Ser Thr Gly Tyr
1      5      10      15
Lys Ser Glu Leu Arg Ser Glu Ala Arg Val Ser Leu Gln Leu Ala Ile
      20      25      30
Pro Leu Val Leu Val Glu Ile Cys Gly Thr Ser Ile Asn Val Val Asp
      35      40      45
Val Val Met Met Gly Leu Leu Gly Thr Gln Val Leu Ala Ala Gly Ala
      50      55      60
Leu Gly Ala Ile Ala Phe Leu Ser Val Ser Asn Thr Cys Tyr Asn Met
65      70      75      80
Leu Leu Ser Gly Val Ala Lys Ala Ser Glu Ala Phe Gly Ala Asn Lys
      85      90      95
Ile Asp Gln Val Ser Arg Ile Ala Ser Gly Gln Ile Trp Leu Ala Leu
      100      105      110
Thr Leu Ser Leu Pro Ala Met Leu Leu Trp Tyr Met Asp Thr Ile
      115      120      125
Leu Val Leu Phe Gly Gln Val Glu Ser Asn Thr Leu Ile Ala Lys Thr
      130      135      140
Tyr Leu His Ser Ile Val Trp Gly Phe Pro Ala Val Gly Ile Leu
145      150      155      160
Ile Leu Arg Gly Ile Ala Ser Ala Val Asn Val Pro Gln Leu Val Thr
      165      170      175
Val Thr Met Leu Val Gly Leu Val Leu Asn Ala Pro Ala Asn Tyr Val
      180      185      190
Leu Met Phe Gly Lys Phe Gly Leu Pro Glu Leu Gly Leu Ala Gly Ile
      195      200      205
Gly Trp Ala Ser Thr Leu Val Phe Trp Ile Ser Phe Leu Val Gly Val
      210      215      220
Val Leu Leu Ile Phe Ser Pro Lys Val Arg Asp Tyr Lys Leu Phe Arg
225      230      235      240
Tyr Leu His Gln Phe Asp Arg Gln Thr Val Val Glu Ile Phe Gln Thr
      245      250      255
Gly Trp Pro Met Gly Phe Leu Leu Gly Val Glu Ser Val Val Leu Ser
      260      265      270
Leu Thr Ala Trp Leu Thr Gly Tyr Leu Gly Thr Val Thr Leu Ala Ala
      275      280      285
His Glu Ile Ala Ile Gln Thr Ala Glu Leu Ala Ile Val Ile Pro Leu
10      290      295      300

```

Gly Ile Gly Asn Val Ala Val Thr Arg Val Gly Gln Thr Ile Gly Glu
 305 310 315 320
 Lys Asn Pro Leu Gly Ala Arg Arg Ala Ala Leu Ile Gly Ile Met Ile
 325 330 335
 Gly Gly Ile Tyr Ala Ser Leu Val Ala Val Ile Phe Trp Leu Phe Pro
 340 345 350
 Tyr Gln Ile Ala Gly Leu Tyr Leu Lys Ile Asn Asp Pro Glu Ser Met
 355 360 365
 Glu Ala Val Lys Thr Ala Thr Asn Phe Leu Phe Leu Ala Gly Leu Phe
 370 375 380
 Gln Phe Phe His Ser Val Gln Ile Ile Val Val Gly Val Leu Ile Gly
 385 390 395 400
 Leu Gln Asp Thr Phe Ile Pro Leu Leu Met Asn Leu Val Gly Trp Gly
 405 410 415
 Leu Gly Leu Ala Val Ser Tyr Tyr Met Gly Ile Ile Leu Cys Trp Gly
 420 425 430
 Gly Met Gly Ile Trp Leu Gly Leu Val Leu Ser Pro Leu Leu Ser Gly
 435 440 445
 Leu Ile Leu Met Val Arg Phe Tyr Gln Glu Ile Ala Asn Arg Ile Ala
 450 455 460
 Asn Ser Asp Asp Gly Gln Glu Ser Ile Ser Ile Asp Asn Val Glu Glu
 465 470 475 480
 Leu Ser

<210> 34

<211> 831

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 34

atgaaaacaa acaaacatat agctatgtgg gcttgtccta gaagtcgttc tactgtaatt 60
 acccgtgctt ttgagaactt agatgggtgt gttgtttatg atgagcctct agaggctcgc 120
 aatgtcttga tgacaactta cacgatgagt aacagtcgta cgtagcaga agaagactta 180
 aagcaattaa tactgcaaaa taatgtagaa acagacctca agaaagttat agaacaattg 240
 actggagatt taccggacgg aaaattattc tcattttcaa aaatgataac aggtgactat 300
 agatctgaat ttggaataga ttgggcaaaa aagctaacta acttcttttt aataaggcat 360
 ccccaagata ttattttttc tttcgatata gcggagagaa agacagggtat cacagaacca 420
 ttcacacaac aaaatcttgg catgaaaaca ctttatgaag ttttccaaca aattgaagtt 480
 attacagggc aaacaccttt agttattcac tcagatgata taattaaaaa ccctccttct 540
 gctttgaaat ggctgtgtaa aaacttaggg cttgcatttg atgaaaagat gctgacatgg 600
 aaagcaaatc tagaagactc caatttaaag tatacaaaat tatatgctaa ttctgcgtct 660
 ggcagttcag aaccttggtt tgaaacttta agatcgacca aaacatttct cgcctatgaa 720
 aagaaggaga aaaaattacc agctcgggta atacctctac tagatgaatc tattccttac 780
 tatgaaaaac tcttacagca ttgtcatatt tttgaatggc cagaacactg a 831

<210> 35

<211> 276

<212> PRT

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 35

Met Lys Thr Asn Lys His Ile Ala Met Trp Ala Cys Pro Arg Ser Arg
 1 5 10 15
 Ser Thr Val Ile Thr Arg Ala Phe Glu Asn Leu Asp Gly Cys Val Val
 20 25 30
 Tyr Asp Glu Pro Leu Glu Ala Pro Asn Val Leu Met Thr Thr Tyr Thr
 35 40 45
 Met Ser Asn Ser Arg Thr Leu Ala Glu Glu Asp Leu Lys Gln Leu Ile
 50 55 60

20

Leu Gln Asn Asn Val Glu Thr Asp Leu Lys Lys Val Ile Glu Gln Leu
 65 70 75 80
 Thr Gly Asp Leu Pro Asp Gly Lys Leu Phe Ser Phe Gln Lys Met Ile
 85 90 95
 Thr Gly Asp Tyr Arg Ser Glu Phe Gly Ile Asp Trp Ala Lys Lys Leu
 100 105 110
 Thr Asn Phe Phe Leu Ile Arg His Pro Gln Asp Ile Ile Phe Ser Phe
 115 120 125
 Asp Ile Ala Glu Arg Lys Thr Gly Ile Thr Glu Pro Phe Thr Gln Gln
 130 135 140
 Asn Leu Gly Met Lys Thr Leu Tyr Glu Val Phe Gln Gln Ile Glu Val
 145 150 155 160
 Ile Thr Gly Gln Thr Pro Leu Val Ile His Ser Asp Asp Ile Ile Lys
 165 170 175
 Asn Pro Pro Ser Ala Leu Lys Trp Leu Cys Lys Asn Leu Gly Leu Ala
 180 185 190
 Phe Asp Glu Lys Met Leu Thr Trp Lys Ala Asn Leu Glu Asp Ser Asn
 195 200 205
 Leu Lys Tyr Thr Lys Leu Tyr Ala Asn Ser Ala Ser Gly Ser Ser Glu
 210 215 220
 Pro Trp Phe Glu Thr Leu Arg Ser Thr Lys Thr Phe Leu Ala Tyr Glu
 225 230 235 240
 Lys Lys Glu Lys Lys Leu Pro Ala Arg Leu Ile Pro Leu Leu Asp Glu
 245 250 255
 Ser Ile Pro Tyr Thr Glu Lys Leu Leu Gln His Cys His Ile Phe Glu
 260 265 270
 Trp Ser Glu His
 275

<210> 36

<211> 774

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 36

ctaaaaaattt ttttctactc ttttcaggat agaattccag tttctagagc cgttgtaacc 60
 gtacatatct tgatagtacg tatcgatgag gtactcattt tcgtggagca ttaaccagct 120
 ttttaactcc gctaatttct gctctccttt ttctattaat tcttgctcat ccaaactcatc 180
 cctgtccaac tcttcctgt ccaactccca catagttttg ttggtatctt cgacaatcaa 240
 gtagtctcca ctttttagac cgttttcgtg aaaatattca actactccca ccgcattagc 300
 atgggcatct tctacgatca accagggatg agcaagccca gaaagcagtt ccgacgacat 360
 tattgcaccc atattgttac aatccccctc taaaaaatga acgcgagagt cagtttttgc 420
 tttctcgtcg agtagggaaa gatcgatc gatacagtag acacaacctt ctatttggaa 480
 cagttctaag tgatcggcta gccaaatcgc gctgccaccg cttaatgctc ctatttcgat 540
 tattgttttc gggcggaagc catacaggag cattgaataa agagctatct cgggtgcaccc 600
 tttcaggaag ggtatccctt tccaagtga ccaatcgcgg tttgccaaga gcgctctcca 660
 agctggcact ggaatagcac atttatcttc tctttcagaa attttggcaa accgattagg 720
 10 tttgaaagg t gcaactttat aggcggcttc ttgaacaaat ttttggaagc tcat 774

<210> 37

<211> 257

<212> PRT

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 37

Met Ser Phe Gln Lys Phe Val Gln Glu Ala Ala Tyr Lys Val Ala Pro
 1 5 10 15
 Phe Lys Pro Asn Arg Phe Ala Lys Ile Ser Glu Arg Glu Asp Lys Cys
 20 25 30

20

Ala Ile Pro Val Pro Ala Trp Arg Ala Leu Leu Ala Asn Arg Asp Leu
35 40 45
Phe Thr Trp Lys Gly Ile Pro Phe Leu Lys Gly Cys Thr Glu Ile Ala
50 55 60
Leu Tyr Ser Met Leu Leu Tyr Glu Leu Arg Pro Lys Thr Ile Ile Glu
65 70 75 80
Ile Gly Ala Leu Ser Gly Gly Ser Ala Ile Trp Leu Ala Asp His Leu
85 90 95
Glu Leu Phe Gln Ile Glu Gly Cys Val Tyr Cys Ile Asp Ile Asp Leu
100 105 110
Ser Leu Leu Asp Glu Lys Ala Lys Thr Asp Ser Arg Val His Phe Leu
115 120 125
Glu Gly Asp Cys Asn Asn Met Gly Ala Ile Met Ser Ser Glu Leu Leu
130 135 140
Ser Gly Leu Ala His Pro Trp Leu Ile Val Glu Asp Ala His Ala Asn
145 150 155 160
Ala Val Gly Val Val Glu Tyr Phe His Glu Asn Gly Leu Lys Ser Gly
165 170 175
Asp Tyr Leu Ile Val Glu Asp Thr Asn Lys Thr Met Trp Glu Leu Asp
180 185 190
Arg Glu Glu Leu Asp Arg Asp Asp Leu Asp Glu Gln Glu Leu Ile Glu
195 200 205
Lys Gly Glu Gln Lys Leu Ala Glu Leu Lys Ser Trp Leu Met Leu His
210 215 220
Glu Asn Glu Tyr Leu Ile Asp Thr Tyr Tyr Gln Asp Met Tyr Gly Tyr
225 230 235 240
Asn Gly Ser Arg Asn Trp Asn Ser Ile Leu Lys Arg Val Glu Lys Asn
245 250 255
Phe

<210> 38

<211> 327

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 38

ttattcaaat agccgtagtt tatgatcggt atccaattcg ctattgtttt ttctgccata 60
tccccaacct aagatgcgac gatattcacc cataatgccca ctgtcaatta aatcatcctc 120
gttgactgca acattggtat gagattgcgg cgcaacatag agcgcattccg caggacaata 180
tgcttcacag atgaaacaag tttgacagtc ttcctgtcgg gcgatcgag gcggttggtt 240
gggaactgca tcaaagacat tggtagggca tacttgagacg caaacattac aattaatata 300
10 gagtttatgg ctgacaagct cgatcat 327

<210> 39

<211> 108

<212> PRT

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 39

Met Ile Glu Leu Val Ser His Lys Leu Cys Ile Asn Cys Asn Val Cys
1 5 10 15
Val Gln Val Cys Pro Thr Asn Val Phe Asp Ala Val Pro Asn Gln Pro
20 25 30
Pro Ala Ile Ala Arg Gln Glu Asp Cys Gln Thr Cys Phe Ile Cys Glu
35 40 45
Ala Tyr Cys Pro Ala Asp Ala Leu Tyr Val Ala Pro Gln Ser His Thr
50 55 60
Asn Val Ala Val Asn Glu Asp Asp Leu Ile Asp Ser Gly Ile Met Gly
65 70 75 80
20 Glu Tyr Arg Arg Ile Leu Gly Trp Gly Tyr Gly Arg Lys Asn Asn Ser
85 90 95
Glu Leu Asp Thr Asp His Lys Leu Arg Leu Phe Glu
100 105

<210> 40

<211> 1653

<212> ADN
<213> *Cylindrospermopsis raciborskii* T3

<400> 40

5

```
ttaagtgggt aatactgggt gtgtagcgct cgcattcctt acccaatccc gtctcaccca 60
aagcctttct aagccgcccc tggcttggtg ataaagctga tttggatcgg tttcaggata 120
gtctatgcga atatgttcgc tacgcgtttc cttgcgatgt aaagcgctaa aatatgcccc 180
tcgtgctaca gacacaagag cagccgctcg acgagaaaat tccagatcgc gcactgtatc 240
ttgtttcggg ttcctcttga cttgctgcca cagcatttct aatttggcga gggaatccaa 300
aagtcctctg tcacagcgca agtaattctt ctctaattggg aacatctcgg cttgtacacc 360
gcggacaact gcctcgctat cgaatgtttc ggaaccaggg tactgggaac gtaatccggc 420
ttgacctgct ggagcgacaa cccgttcctg gacatgagcg cccaaactct tggcaaaggc 480
ggctgcacct tcccttgccc attgtcctgt agagattgcc caagcagcat taggaccatc 540
accccagaaa gctatcccag ctaaaaactc ccgcgatgct gcatctccgg cggcatacag 600
tccaggaact tttgtaccac aactatcatt cacaatccga attccacctg taccacggac 660
tgtaccttct aaaaccagtg ttacaggtac tcgttctgta taagggtcaa tgccagcttt 720
tttatagggt agaaaggcga tgaagtgaga cttttcaacc aatgcttgga tttcaggtgt 780
ggctcgatcc aaacgagcat aaacgggacc tttcaggagg gcattgggca ggaacgatgg 840
atcgcgacga ccattgatat agccaccaag atcgttacct gcctcatcgg tgtaactagc 900
ccagtaaaag ggagcagccc ttgtcactgt ggcattgaaa gcggtcgaga tggatatagt 960
actggaagct tccatactgg agagttcgcc gccagcttcc accgccatca gcagtccatc 1020
gcctgtattg gtattgcaac ctaaagcttt acttaggaat gcacaaccgc cattcgctag 1080
aactactgca ccagcgcgaa cgggtatagg gcgatgattt tgctctgta cacctctagc 1140
tccagccacg gagccgtcct gggctaataa cagtcttaga gccggacttt ggtcgaaaat 1200
ttgcacaccc acacgcaaca ggttcttgcg aagtaccgcg atatatcccg gaccataata 1260
actctggcgc acggattccc cattttcttt ggggaaacga tagccccaat cttccactaa 1320
gggcaaaact agccaagctt tttcaattac acgttcaatc caacgtaagt tagcgagggt 1380
atttcctttg ctgtaacatt cggatacatc tttctcccaa ttctctggag aagggtgccat 1440
gacgctattg ccactggcag cagctgcacc gctcgtacct agaaaacctt tatcaacaat 1500
gatgactttg acaccttggg ctccagccgc ccatgctgcc catgcgggcg caggaccacc 1560
accaattacc agcacgtcag cagttaattg tagttcagtg ccgctatagg ctgtaagcaa 1620
ttgcttttcc tccttgttta aagtcaagtt cat 1653
```

<210> 41
<211> 550
<212> PRT
<213> *Cylindrospermopsis raciborskii* T3

<400> 41

10

```
Met Asn Leu Thr Leu Asn Lys Glu Glu Lys Gln Leu Leu Thr Ala Tyr
1      5      10      15
Ser Gly Thr Glu Leu Gln Leu Thr Ala Asp Val Leu Val Ile Gly Gly
20     25     30
Gly Pro Ala Ala Ala Trp Ala Ala Trp Ala Ala Gly Ala Gln Gly Val
35     40     45
Lys Val Ile Ile Val Asp Lys Gly Phe Leu Gly Thr Ser Gly Ala Ala
50     55     60
Ala Ala Ser Gly Asn Ser Val Met Ala Pro Ser Pro Glu Asn Trp Glu
65     70     75     80
Lys Asp Val Ser Glu Cys Tyr Ser Lys Gly Asn Asn Leu Ala Asn Leu
85     90     95
15 Arg Trp Ile Glu Arg Val Ile Glu Lys Ala Trp Leu Ser Leu Pro Leu
```

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

<210> 42

<211> 750

<212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 42

```

ttaattatct tctgcagtcg gtcgaatcaa aatttcattt acatttacat gatcgggttg 60
tgtcactgca taaattatag ctcttgcaat atcctcactt tgtaaagggtg ttattgtact 120
aagttgttct ttactaagct gtttcgtgat cgggtcagaa attaagtcac taaatggcgt 180
atcgactaaa cctgggtcaa tgatggtaac gcgaatgttg tctaaagata cctcctggcg 240
taatgcttct gaaagagcat tgacgcctga ttgggcagca ctataaacga ccgcaccgga 300
ctgcgctatc ctgccatcga cagaagatat attgactata tgaccggatt tttgggcctt 360
cagaagaggc aaaactgcgt ggatagcata taaaactccc agaacattca catcgaatgc 420
tcgcctccag tctgcgggat ttccagtatc aattgcacca aacacaccaa ttcttgcat 480
attcaccaaa atatctacat gtccctagctc aaccttggtc ttttgacta gatgatttac 540
ttgagattcg tctgtaatat ctgtaacaat aggcgaatgct tgaccaccac tggcttcaat 600
ccgttttgct agtgcacgca aaagctcagc acgtcttgcg gcgatcgcaa cttttgcccc 660
ctccgcagct aaagcaaatg ctgtagcctc tccaatccca gaggaagctc cagtaataat 720
cgccactttt ccatccaatt tacctgccat 750

```

<210> 43

<211> 249

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 43

```

Met Ala Gly Lys Leu Asp Gly Lys Val Ala Ile Ile Thr Gly Ala Ser
1      5      10      15
Ser Gly Ile Gly Glu Ala Thr Ala Phe Ala Leu Ala Ala Glu Gly Ala
      20      25      30
Lys Val Ala Ile Ala Ala Arg Arg Ala Glu Leu Leu His Ala Leu Ala
      35      40      45
Lys Arg Ile Glu Ala Ser Gly Gly Gln Ala Leu Pro Ile Val Thr Asp
      50      55      60
Ile Thr Asp Glu Ser Gln Val Asn His Leu Val Gln Lys Thr Lys Val
      65      70      75      80
Glu Leu Gly His Val Asp Ile Leu Val Asn Asn Ala Gly Ile Gly Val
      85      90      95
Phe Gly Ala Ile Asp Thr Gly Asn Pro Ala Asp Trp Arg Arg Ala Phe
      100     105     110
Asp Val Asn Val Leu Gly Val Leu Tyr Ala Ile His Ala Val Leu Pro
      115     120     125
Leu Leu Lys Ala Gln Lys Ser Gly His Ile Val Asn Ile Ser Ser Val
      130     135     140
Asp Gly Arg Ile Ala Gln Ser Gly Ala Val Val Tyr Ser Ala Ala Lys
      145     150     155     160
Ser Gly Val Asn Ala Leu Ser Glu Ala Leu Arg Gln Glu Val Ser Leu
      165     170     175
Asp Asn Ile Arg Val Thr Ile Ile Glu Pro Gly Leu Val Asp Thr Pro
      180     185     190
Phe Asn Asp Leu Ile Ser Asp Pro Ile Thr Lys Gln Leu Ser Lys Glu
      195     200     205
Gln Leu Ser Thr Ile Thr Pro Leu Gln Ser Glu Asp Ile Ala Arg Ala
      210     215     220
Ile Ile Tyr Ala Val Thr Gln Pro Asp His Val Asn Val Asn Glu Ile
      225     230     235     240
Leu Ile Arg Pro Thr Ala Glu Asp Asn
      245

```

10

<210> 44

<211> 1005

<212> ADN

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 44

```

ttaacaaacc ccataagtaa cacctagttg ctttagccat cgacgatagg caagtgtgca 60
tctatctgat ggtacgtgga tttcgtgtga aaacaattgt gtatttatct gctttggagt 120
taacagtggt aaacgtaccg gctgtgtgac atgtaagatc cgaatatctt gttctattgt 180
ttcgtcatat tcagtttagca tctttgactc taacgtttca taccogttcc acattatcaa 240
catacgcaat acactatttt cctcatcaat cgggtgtgatc gtcattaaat ccacaatcct 300
catttcaggg gattctgaaa cgcagtattg acataaagga tgactaagcc tgaaccaatt 360
aaccgaagag tcattcttga tatggctgac aatccttgat gtctggaatt gatacttacc 420
catagtaagg ccatctttat ctaatttcac ctcaaattct tccacttttg tataattgcg 480
atcacctaac caaccgtcat ggataaaaagg aaaatgagac acgtctaagg aattatccat 540
cacacgaaac gcactagctt taatcaagta agacttggtg taagtcttgt gataattcgg 600
atcatcccat tcaggaaatg aaggtatatc attaacagga tcgccaagc acaccacac 660
taagccatag cgctcctggg agtgatatgt cctggcttca gcacttgccg gtggtaccat 720
gccaggggtg gctgggatct gtatgcattt accagcctca ttgtatctcc atccgtgata 780
cggacaaact aaagtattat tcgtaatttc tcccatagac agaggaacac ctcggtgggg 840
gcagtagtca agccatacct gtatgggtga attttgttca taactgcgcc ataataccaa 900
cttactctcc aacaaacgag atctggtgat acttccaggt ttacagtctt ctacattggc 960
gactacgtgc cagttattga ttaagattgg gtcggtagtt gtcatt 1005

```

<210> 45

<211> 334

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 45

```

Met Thr Thr Thr Asp Pro Ile Leu Ile Asn Asn Trp His Val Val Ala
1      5      10      15
Asn Val Glu Asp Cys Lys Pro Gly Ser Ile Thr Arg Ser Arg Leu Leu
20     25     30
Gly Val Lys Leu Val Leu Trp Arg Ser Tyr Glu Gln Asn Ser Pro Ile
35     40     45
Gln Val Trp Leu Asp Tyr Cys Pro His Arg Gly Val Pro Leu Ser Met
50     55     60
Gly Glu Ile Thr Asn Asn Thr Leu Val Cys Pro Tyr His Gly Trp Arg
65     70     75     80
Tyr Asn Glu Ala Gly Lys Cys Ile Gln Ile Pro Ala His Pro Gly Met
85     90     95
Val Pro Pro Ala Ser Ala Glu Ala Arg Thr Tyr His Ser Gln Glu Arg
100    105    110
Tyr Gly Leu Val Trp Val Cys Leu Gly Asp Pro Val Asn Asp Ile Pro
115    120    125
Ser Phe Pro Glu Trp Asp Asp Pro Asn Tyr His Lys Thr Tyr Thr Lys
130    135    140
Ser Tyr Leu Ile Lys Ala Ser Ala Phe Arg Val Met Asp Asn Ser Leu
145    150    155    160
Asp Val Ser His Phe Pro Phe Ile His Asp Gly Trp Leu Gly Asp Arg
165    170    175
Asn Tyr Thr Lys Val Glu Glu Phe Glu Val Lys Leu Asp Lys Asp Gly
180    185    190
Leu Thr Met Gly Lys Tyr Gln Phe Gln Thr Ser Arg Ile Val Ser His
195    200    205
Ile Glu Asp Asp Ser Trp Val Asn Trp Phe Arg Leu Ser His Pro Leu
210    215    220

Cys Gln Tyr Cys Val Ser Glu Ser Pro Glu Met Arg Ile Val Asp Leu
225    230    235    240
Met Thr Ile Thr Pro Ile Asp Glu Glu Asn Ser Val Leu Arg Met Leu
245    250    255
Ile Met Trp Asn Gly Tyr Glu Thr Leu Glu Ser Lys Met Leu Thr Glu
260    265    270
Tyr Asp Glu Thr Ile Glu Gln Asp Ile Arg Ile Leu His Ala Gln Gln
275    280    285
Pro Val Arg Leu Pro Leu Leu Thr Pro Lys Gln Ile Asn Thr Gln Leu
290    295    300
Phe Ser His Glu Ile His Val Pro Ser Asp Arg Cys Thr Leu Ala Tyr
305    310    315    320
Arg Arg Trp Leu Lys Gln Leu Gly Val Thr Tyr Gly Val Cys
325    330

```

10

<210> 46

15 <211> 726

<212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 46

5

```

ctaaattatc cttttcaagg catccaccaa cagtgggttg atgttggttt ttgtaaaaat    60
cagagtttagc atcctgtaat cggtaattga agtggtggca gctgcggtat gccatacagt    120
tggtgtataa aacattgctg cccctcctgg aagtgaaga catatctctg catttagtga    180
attggcagaa gatgaatcta atgagtgttc ccattggtgg ctacttggtg taactcgcac    240
tgtaccata gtattatctg tatcctgtaa gtatatagtt atgaatacca tggcttgatt    300
ggctactgga accaacaacc gaagcgcgtc gtcatttaac tcgttttttg acatggatgc    360
aagtgcgttc aatacttcaa ctacatatcc atggctctga tgccaagcaa tgtatcctgt    420
acctgcacga attatggcta gatcgggtgat caataggaag atatcagacc caattagagc    480
ctgtactggt cccatcacag ttggaagctc taaaagcctc tgaattatct ttgtatacct    540
aactggatct gggatagtat gctcagacca ccactcatag tcacccgccca atactcccc    600
acgtttttgt tcggaataaa gttctacttc atgccgtatt tcttcaatta acgcttttgg    660
tacagtttct tcaactgtga aataaccatc atttgtgtaa gcttggtttt gttccgctgt    720
gagcat

```

<210> 47

<211> 241

10

<212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 47

15

```

Met Leu Thr Ala Glu Gln Lys Gln Ala Tyr Thr Asn Asp Gly Tyr Phe
1      5      10      15
Thr Val Glu Glu Ala Val Pro Lys Ala Leu Ile Glu Glu Ile Arg His
20     25     30
Glu Val Glu Leu Ile Thr Glu Gln Lys Arg Gly Gly Val Leu Ala Gly
35     40     45
Asp Tyr Glu Trp Trp Ser Glu His Thr Ile Pro Asp Pro Val Arg Tyr
50     55     60
Gln Lys Ile Ile Gln Arg Leu Leu Glu Leu Pro Thr Val Met Gly Pro
65     70     75     80
Val Gln Ala Leu Ile Gly Ser Asp Ile Phe Leu Leu Ile Thr Asp Leu
85     90     95
Ala Ile Ile Arg Ala Gly Thr Gly Tyr Ile Ala Trp His Gln Asp His
100    105    110
Gly Tyr Val Val Glu Val Leu Asn Ala Leu Ala Ser Met Ser Lys Asn
115    120    125
Glu Leu Asn Asp Asp Ala Leu Arg Leu Leu Val Pro Val Ala Asn Gln
130    135    140

Ala Met Val Phe Ile Thr Ile Tyr Leu Gln Asp Thr Asp Asn Thr Met
145    150    155    160
Gly Thr Met Arg Val Ile Pro Ser Ser His Gln Trp Glu His Ser Leu
165    170    175
Asp Ser Ser Ser Ala Asn Ser Leu Asn Ala Glu Ile Cys Leu Ser Leu
180    185    190
Pro Gly Gly Ala Ala Met Phe Tyr Thr Pro Thr Val Trp His Thr Ala
195    200    205
Ala Ala Asn Thr Ser Ile Thr Asp Tyr Arg Met Leu Thr Leu Ile Phe
210    215    220
Thr Lys Asn Asn Ile Lys Pro Leu Leu Val Asp Ala Leu Lys Arg Ile
225    230    235    240
Ile

```

<210> 48

20

<211> 576

<212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 48

25

```

tcaatgggta gtaggaatta tcctatagct gttctttctc tggatagaag aaagggttg 60
agaagctcgc tccgacttca tttcagccaa tttttctgca gaccaatact gaaaatatcc 120
caatcttaat aattcatcac tagcctcttg taactggctg aatgactgta ctgatgctaa 180
aacatactta gggtaggta tgattacgtt attcacattc tccgcgtcat caccaacata 240
ttgtttgtct ggatgcgac ctaaagctac caaatcgat tctggtaata cataattcgc 300
cttggaatg tacctttcca acctctgtgc atctagggtt tgagggtcgc agccaaaaat 360
caccatttca aagtcattat tccatgttct tatctgttcc attagaagct ctggcagttc 420
agggtccatga aaccaacgaa cactaacacg gttattttaac caagctgcct tcgcgtaagg 480
acagggtgga aaatttctctg ttagaggatt gggaatgctg acaacattga taatccaatc 540
ctctatttct tggcgaaatt gttcgatatt tatcat 576

```

<210> 49

<211> 191

<212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 49

```

Met Ile Asn Ile Glu Gln Phe Arg Gln Glu Ile Glu Asp Trp Ile Ile
1           5           10          15
Asn Val Val Ser Ile Pro Asn Pro Leu Thr Gly Asn Phe Pro Pro Cys
20          25          30
Pro Tyr Ala Lys Ala Ala Trp Leu Asn Asn Arg Val Ser Val Arg Trp
35          40          45
Phe His Gly Pro Glu Leu Pro Glu Leu Leu Met Glu Gln Ile Arg Thr
50          55          60
Trp Asn Asn Asp Phe Glu Met Val Ile Phe Gly Cys Asp Pro Gln Asn
65          70          75          80
Leu Asp Ala Gln Arg Leu Glu Arg Tyr Ile Thr Lys Ala Asn Tyr Val
85          90          95
Leu Pro Glu Tyr Asp Leu Val Ala Leu Gly Ser His Pro Asp Lys Gln
100         105         110
Tyr Val Gly Asp Asp Ala Glu Asn Val Asn Asn Val Ile Ile Thr His
115         120         125
Pro Lys Tyr Val Leu Ala Ser Val Gln Ser Phe Ser Gln Leu Gln Glu
130         135         140
Ala Ser Asp Glu Leu Leu Arg Leu Gly Tyr Phe Gln Tyr Trp Ser Ala
145         150         155         160
Glu Lys Leu Ala Glu Met Lys Ser Glu Arg Ala Ser His Asn Leu Ser
165         170         175
Ser Ile Gln Arg Lys Asn Ser Tyr Arg Ile Ile Pro Thr Asn His
180         185         190

```

<210> 50

<211> 777

<212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 50

```

ttaatctagg tcatagtata accatatatt aggctcgatg tatattccca tattgttggg 60
atagtcaatt ttgacaggta ctaagccttt gggaataata tagtcaccag tttctggaaa 120
acgcatacca actctatctt cccaaccgtc aatagttatca ttaattgttg tggatttaaa 180
acagatcccc gcaatttttag ccccatgttt gacattaact cgtaaccaag ggtcaaatat 240
aagaccattt ttatctcgcc aggtaatata ccgctctatg ggtataagtg ggtaaagata 300
ttttaggctt ggacgtgcag ccatgatcaa agaattaaga ccgtgggtatt gagcaagttc 360
tttcatgtat ccaatcagat actgactcaa gtttttgctt tgataactctg gtaggattga 420
aatcgatact acacataacg cattaggcag gcggttctgt tctcgggtctt caagccactt 480
gggtaaaagg cagtcacaac cttcgtccgg taactcatca aaacggcttt cataagttaa 540
agggatacac tttccttgcg ctatcataag ctgtgtggta gcttctacta acccaactg 600
gaattctgga taaatttcaa atagagctaa ggaagctgga tctgcccaga catcatgtat 660
caaaaatttt ggggtatgctt gatcaaagac actcatcgtc ctttccacaa aatcagaagt 720
ttcttttggg gttacaaagc tatactctaa attatgctgt acaatttgaa tggtcatt 777

```

<210> 51

<211> 258

<212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 51

Met Thr Ile Gln Ile Val Gln His Asn Leu Glu Tyr Ser Phe Val Thr
1 5 10 15
Pro Lys Glu Thr Ser Asp Phe Val Glu Arg Thr Met Ser Val Phe Asp
20 25 30
Gln Ala Tyr Pro Lys Phe Leu Ile His Asp Val Trp Ala Asp Pro Ala
35 40 45
Ser Leu Ala Leu Phe Glu Ile Tyr Pro Glu Phe Gln Phe Gly Leu Val
50 55 60
Glu Ala Thr Thr Gln Leu Met Ile Ala Gln Gly Asn Cys Ile Pro Leu
65 70 75 80
Thr Tyr Glu Ser Arg Phe Asp Glu Leu Pro Asp Glu Gly Cys Asp Trp
85 90 95
Ala Leu Ala Lys Trp Leu Glu Asp Arg Glu Gln Asn Arg Leu Pro Asn
100 105 110
Ala Leu Cys Val Val Ser Ile Ser Ile Leu Pro Glu Tyr Gln Gly Lys
115 120 125
Asn Leu Ser Gln Tyr Leu Ile Gly Tyr Met Lys Glu Leu Ala Gln Tyr
130 135 140
His Gly Leu Asn Ser Leu Ile Met Ala Ala Arg Pro Ser Leu Lys Tyr
145 150 155 160
Leu Tyr Pro Leu Ile Pro Ile Glu Arg Tyr Ile Thr Trp Arg Asp Lys
165 170 175
Asn Gly Leu Ile Phe Asp Pro Trp Leu Arg Val Asn Val Lys His Gly
180 185 190
Ala Lys Ile Ala Gly Ile Cys Phe Lys Ser Thr Thr Ile Asn Asp Thr
195 200 205
Ile Asp Gly Trp Glu Asp Arg Val Gly Met Arg Phe Pro Glu Thr Gly
210 215 220
Asp Tyr Ile Ile Pro Lys Gly Leu Val Pro Val Lys Ile Asp Tyr Pro
225 230 235 240
Asn Asn Met Gly Ile Tyr Ile Glu Pro Asn Ile Trp Leu Tyr Tyr Asp
245 250 255
Leu Asp

<210> 52

<211> 777

<212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 52

ctaatacctta aatttataact ggaagtcaaa tgagatctca ctatcggttat tatctggaag 60
tacttgact gtcaattcat taccgacttt cccattccca ggcataatta ataagttagg 120
gtgaggtgga atgccgtcgt actgtcggac gcggcgaaaa atgctcgaat tctcgccacc 180
atgtttattc aagaggactt caactggtgt gatgacaaaa gtcattcctg acccaagggtg 240
gcgcgatcgc cgcttttgat ttgctggagt ggaaacacta acaaataagg cacaccctcc 300
tagagaataa gaccagttag cagactgcgg atcggcagac caatggcagg gacaagacac 360
cgcatcaagg ctatgtaacg cattcaaaaa atcaaatgct tgacctgcat attcctctac 420
tgtaagaact gttggttcag gtgggaaaaa gatgacaagt gtcagaagat ccgcattttc 480
gtgctgaagc aattcgtttt cattaacttc atcaatgtat ttgtagatac cctcaagcgt 540
atgctcaacc aagatcgggt cagttaaaga tgagactatc aggtatctaa tcattccctt 600
ctgttccccg atagttcccc agaagcaagg gaaggcagaa tcgctgattg tttcaacaaa 660
tgttgagtag ctagtgcgta cccaagcagg aaggcactcc tctagaagag aggattccat 720
ctggcttttg ttccagattg gtgtaactcc gtcaggacat aaattcttga ttaccat 777

<210> 53

<211> 258

<212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 53

ES 2 525 045 T3

Met Val Ile Lys Asn Leu Cys Pro Asp Gly Val Thr Pro Ile Trp Asn
 1 5 10 15
 Lys Ser Gln Met Glu Ser Ser Leu Leu Glu Glu Cys Leu Pro Ala Trp
 20 25 30
 Val Arg Thr Ser Tyr Ser Thr Phe Val Glu Thr Ile Ser Asp Ser Ala
 35 40 45
 Phe Pro Cys Phe Trp Gly Thr Ile Gly Glu Gln Lys Gly Met Ile Arg
 50 55 60
 Tyr Leu Ile Val Ser Ser Leu Thr Asp Pro Ile Leu Val Glu His Thr
 65 70 75 80
 Leu Glu Gly Ile Tyr Lys Tyr Ile Asp Glu Val Asn Glu Asn Glu Leu
 85 90 95
 Leu Gln His Glu Asn Ala Asp Leu Leu Thr Leu Val Ile Phe Phe Pro
 100 105 110
 Pro Glu Pro Thr Val Leu Thr Val Glu Glu Tyr Ala Gly Gln Ala Phe
 115 120 125
 Asp Phe Leu Asn Ala Leu His Ser Leu Asp Ala Val Ser Cys Pro Cys
 130 135 140
 His Trp Ser Ala Asp Pro Gln Ser Ala Asn Trp Ser Tyr Ser Leu Gly
 145 150 155 160
 Gly Cys Ala Leu Phe Val Ser Val Ser Thr Pro Ala Asn Gln Lys Arg
 165 170 175
 Arg Ser Arg His Leu Gly Ser Gly Met Thr Phe Val Ile Thr Pro Val
 180 185 190
 Glu Val Leu Leu Asn Lys His Gly Gly Glu Asn Ser Ser Ile Phe Arg
 195 200 205
 Arg Val Arg Gln Tyr Asp Gly Ile Pro Pro His Pro Asn Leu Leu Ile
 210 215 220

Met Pro Gly Asn Gly Lys Val Gly Asn Glu Leu Thr Val Gln Val Leu
 225 230 235 240
 Pro Asp Asn Asn Asp Ser Glu Ile Ser Phe Asp Phe Gln Tyr Lys Phe
 245 250 255

Lys Asp

- 5 <210> 54
 <211> 1227
 <212> ADN
 <213> Cylindrospermopsis raciborskii T3

- 10 <400> 54

ctatatctta ttttttgga gtcctgaaa attattcaac aagatcgaga cggtgttgtt 60
 gccagaattt gtgacagcca ggtcaagctt gctgtcgccg ttgaaatccg caattgctat 120
 agattcagga ttagtaccga ctggaaagtt agtagctatg ccaaaagacc cattaccatt 180
 tcctggtaag accgagacgt tattgtactt ataatttcta acagccaggt caagtttact 240
 gtcgccattc acatctctaa tcgctacaga gtagggatta gtaccggctg gaaagttagt 300
 ggctgcgcca aaagacccat taccatttcc cagtaagacc gagacggtat tgctgctagt 360
 atttgaaca gccaggtcaa gcttgcgtgc gccatttaca tccccagttg ctacaaatat 420
 gggattagta ccgactggaa agttagtggc tgcgcaaaa gaccattac catttccag 480
 taagaccgag acgttattgc tgaccaatt tgtaatagca agtcgagct tactgtcgt 540
 attaaaaatcc gcaatcgcta cggaaatcga ataagtatcg acagggaagc tgctggctgc 600
 gccaaaagac ccattaccat ttcccagtaa aaccaagacc ttattgtcga accaatttgt 660
 aaaagcaagg tcaagctcac tatcgttatt cacatctcca atggctacag aataagggtt 720
 agtaccact gaaaagttag tggctgcgcc aaaagaccca ttaccatttc ctagtaagac 780
 cgagacgtta ttgtactaa aatttgcaac agccaggtca agcttgcgtg cgccatttac 840
 atccccagtc actacaaaga cgggattagt accgactgga aagttagtgg ctgcgcaaaa 900
 agaccatta ccatttccca gtaagaccga gacgttattg tcgaaccaat ttgtaacagc 960
 caggtcgagc ttactatcgc tattgaaatc cccaactgct acagagtcag catcaagacc 1020
 agttgggaag ttaatagcag tagcataact actcctgtgg gcaaatctca ctctacgga 1080
 caaattaacc ggaacactaa attgcccaga aagcttttca ttcttcagat aatagtcagt 1140
 tataattgct aatgcaacag gagttatata taaaaatgta ctaacagata atatccccgc 1200
 tataattagt aaagtgagcc ttttcac 1227

- 15 <210> 55
 <211> 408
 <212> PRT
 <213> Cylindrospermopsis raciborskii T3

<400> 55

Met	Lys	Arg	Leu	Thr	Leu	Leu	Ile	Ile	Ala	Gly	Ile	Leu	Ser	Val	Ser
1				5					10						15
Thr	Phe	Leu	Cys	Ile	Thr	Pro	Val	Ala	Leu	Ala	Asn	Ile	Thr	Asp	Tyr
		20						25					30		
Tyr	Leu	Lys	Asn	Glu	Lys	Leu	Ser	Gly	Gln	Phe	Ser	Val	Pro	Val	Asn
		35					40					45			
Leu	Ser	Val	Gly	Val	Arg	Phe	Ala	His	Arg	Ser	Ser	Tyr	Ala	Thr	Ala
	50					55					60				
Ile	Asn	Phe	Pro	Thr	Gly	Leu	Asp	Ala	Asp	Ser	Val	Ala	Val	Gly	Asp
65					70					75					80
Phe	Asn	Ser	Asp	Ser	Lys	Leu	Asp	Leu	Ala	Val	Thr	Asn	Trp	Phe	Asp
				85					90					95	
Asn	Asn	Val	Ser	Val	Leu	Leu	Gly	Asn	Gly	Asn	Gly	Ser	Phe	Gly	Ala
			100					105					110		
Ala	Thr	Asn	Phe	Pro	Val	Gly	Thr	Asn	Pro	Val	Phe	Val	Val	Thr	Gly
		115					120					125			
Asp	Val	Asn	Gly	Asp	Ser	Lys	Leu	Asp	Leu	Ala	Val	Ala	Asn	Phe	Ser
	130					135					140				
Ser	Asn	Asn	Val	Ser	Val	Leu	Leu	Gly	Asn	Gly	Asn	Gly	Ser	Phe	Gly
145					150					155					160
Ala	Ala	Thr	Asn	Phe	Ser	Val	Gly	Thr	Asn	Pro	Tyr	Ser	Val	Ala	Ile
				165					170					175	
Gly	Asp	Val	Asn	Asn	Asp	Ser	Glu	Leu	Asp	Leu	Ala	Phe	Thr	Asn	Trp
			180					185					190		
Phe	Asp	Asn	Lys	Val	Leu	Val	Leu	Leu	Gly	Asn	Gly	Asn	Gly	Ser	Phe
		195					200					205			
Gly	Ala	Ala	Ser	Ser	Phe	Pro	Val	Asp	Thr	Tyr	Ser	Ile	Ser	Val	Ala
	210					215					220				
Ile	Ala	Asp	Phe	Asn	Ser	Asp	Ser	Lys	Leu	Asp	Leu	Ala	Ile	Thr	Asn
225					230					235				240	
Trp	Val	Ser	Asn	Asn	Val	Ser	Val	Leu	Leu	Gly	Asn	Gly	Asn	Gly	Ser
				245					250					255	
Phe	Gly	Ala	Ala	Thr	Asn	Phe	Pro	Val	Gly	Thr	Asn	Pro	Ile	Phe	Val
			260					265					270		
Ala	Thr	Gly	Asp	Val	Asn	Gly	Asp	Ser	Lys	Leu	Asp	Leu	Ala	Val	Ala
		275					280					285			
Asn	Thr	Ser	Ser	Asn	Asn	Val	Ser	Val	Leu	Leu	Gly	Asn	Gly	Asn	Gly
	290					295					300				
Ser	Phe	Gly	Ala	Ala	Thr	Asn	Phe	Pro	Ala	Gly	Thr	Asn	Pro	Tyr	Ser
305					310					315					320
Val	Ala	Ile	Arg	Asp	Val	Asn	Gly	Asp	Ser	Lys	Leu	Asp	Leu	Ala	Val
				325					330					335	
Thr	Asn	Tyr	Ser	Ser	Asn	Asn	Val	Ser	Val	Leu	Pro	Gly	Asn	Gly	Asn

5

<210> 56
<211> 603
<212> ADN
<213> *Cylindrospermopsis raciborskii* T3

10

<400> 56

```

ctattgtttg aaaattgtga atttgttttc cacgtatttg agtagttgtt ctaggctttc 60
ctcgacggtg agttcggatg tttccacca taaatctggg ctattgggtg gttcataagg 120
ggcgctgatt ccgtaaatc catctatttc cccactgctg gcttttagat aaagaccttt 180
cggatcacgc tgctcacaaa gttccagtgg agttgcaatg tatacttcat gaaatagatc 240
tccagctagt ctacgcacct gttctcggtc attcctgtag ggtgagatga aggcagtgat 300
cactaggcat cctgactccg caaagagttt ggcaacctca cccaaacgac ggatattttc 360
tgagcgaatc ctacgagaaa atcctaaatc ggaacacagt ccatgacgaa cactatcacc 420
atctaaaaca aaggtagacc atcctttctc gaacaaagtc tgctctaatt ttaaagccaa 480
tggtgtttta ccagccccgg acagtccagt aaaccataga atcccgtttt tatgaccatt 540
ctttagataa cgatcatatg gagatataag atgttttgta tagtgaatat tagttgattt 600
cat 603

```

<210> 57

<211> 200

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 57

```

Met Lys Ser Thr Asn Ile His Tyr Thr Lys His Leu Ile Ser Pro Tyr
1 5 10 15
Asp Arg Tyr Leu Lys Asn Gly His Lys Ser Gly Ile Leu Trp Phe Thr
20 25 30
Gly Leu Ser Gly Ala Gly Lys Thr Thr Leu Ala Leu Lys Leu Glu Gln
35 40 45
Thr Leu Phe Glu Lys Gly Trp Ser Thr Phe Val Leu Asp Gly Asp Ser
50 55 60
Val Arg His Gly Leu Cys Ser Asp Leu Gly Phe Ser Ala Ser Asp Arg
65 70 75 80
Ser Glu Asn Ile Arg Arg Leu Gly Glu Val Ala Lys Leu Phe Ala Glu
85 90 95
Ser Gly Cys Leu Val Ile Thr Ala Phe Ile Ser Pro Tyr Arg Asn Asp
100 105 110
Arg Glu Gln Val Arg Arg Leu Ala Gly Asp Leu Phe His Glu Val Tyr
115 120 125
Ile Ala Thr Pro Leu Glu Leu Cys Glu Gln Arg Asp Pro Lys Gly Leu
130 135 140
Tyr Leu Lys Ala Arg Ser Gly Glu Ile Asp Gly Phe Thr Gly Ile Ser
145 150 155 160
Ala Pro Tyr Glu Pro Pro Asn Ser Pro Asp Leu Trp Val Glu Thr Ser
165 170 175
Glu Leu Thr Val Glu Glu Ser Leu Glu Gln Leu Leu Lys Tyr Val Glu
180 185 190
Asn Lys Phe Thr Ile Phe Lys Gln
195 200

```

10

<210> 58

<211> 1350

<212> ADN

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 58

ES 2 525 045 T3

ttaagaaaaa	attatittcaa	actcgctcgc	caaacgctcc	ataatcaaat	taatttcaga	60
cgaaaaagga	cagtaatatg	gtagctctac	caacaccctt	cttgcggaag	ctgtcacctt	120
cgctgctatt	ttgataatcg	tttcccttaa	cctaggaacc	tgggcttttag	ccagttttgt	180
tccctgtgct	gcttgccgaa	ttcccaacat	taaaatgtaa	gctgcttgag	ataaaaaata	240
ccgaaactga	ttgacaataa	atttctcaca	gctgagtcta	tctgatttta	tcccagttt	300
taatttcctta	attctatgct	ctgaagtagc	tcctctttga	acataaaaatt	tatcgtataa	360
atcctgagct	tctgtttcca	agctagtaat	tataaatcta	ggattgggtc	ctttttctag	420
ccattctgct	ttcataatta	ctcgccgagg	ttctgaccaa	ctccgagctg	cgtaatacac	480
atcatcaaat	aaacgaactt	tttctcctgt	gcgacaatat	tccagtcctg	ctcgggtcaag	540
aaggtaatta	atttttcggt	ttaagacatc	attattgctg	aatccaaaaa	catatccaac	600
cccgtttttt	tcacaaacct	caatgatttc	tggtaacgag	aaacccccgt	ctcccctcag	660
aacaattcta	atttcaggta	aggctctttt	gattcgcaaa	aataaccatt	ttagaatgcc	720
agctactcct	ttaccagagt	gagaatttcc	cgcccttagt	tgtagaacta	atggataacc	780
actggaagct	tcattaatca	gaactggaaa	gtagatatca	tgcttatggt	aaccattaaa	840
taagctcagt	tgttgatgac	catgagttag	agcatcccac	gcatctatgt	ccaggacaat	900
ctcttttgat	tcccgaggat	aggattctag	gaatttatca	acaaataacc	gacgaatttg	960
tttgatatct	ttttgagtca	cctgattttc	taaacgactc	atagttgggt	gactagctaa	1020
taagttttct	cctactgtgg	gaacttgatt	acaaactagc	ttaaaaattg	gatcttggcg	1080
caatttatta	ctatcggttc	tatcttcata	gccagcaatt	atttgataaa	ttcgttggct	1140
aattaattga	gaaagagaat	gtttgacttt	agtttggtcc	cgattatccg	tcaaacaatc	1200
tgccatatct	tgacaaattt	ttaccttttc	ttctacttgt	cgtgccagaa	taattccgcc	1260
atcactactt	aaactcatat	cagaaaaagt	cagatctaaa	gtttttttat	cgaagaaatt	1320
ttaagataat	cttgaggaag	atttagtcat				1350

<210> 59

<211> 449

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 59

```

Met Thr Lys Ser Ser Ser Arg Leu Ser Leu Asn Phe Phe Asp Lys Lys
1      5      10      15
Thr Leu Asp Leu Thr Phe Ser Asp Met Ser Leu Ser Ser Asp Gly Gly
      20      25      30
Ile Ile Leu Ala Arg Gln Val Glu Glu Lys Val Lys Ile Cys Gln Asp
      35      40      45
Met Ala Asp Cys Leu Thr Asp Asn Arg Asp Gln Thr Lys Val Lys His
      50      55      60
Ser Leu Ser Gln Leu Ile Ser Gln Arg Ile Tyr Gln Ile Ile Ala Gly
      65      70      75      80
Tyr Glu Asp Ser Asn Asp Ser Asn Lys Leu Arg Gln Asp Pro Ile Phe
      85      90      95
Lys Leu Val Cys Asn Gln Val Pro Thr Val Gly Glu Asn Leu Leu Ala
      100      105      110
Ser Gln Pro Thr Met Ser Arg Leu Glu Asn Gln Val Thr Gln Lys Asp
      115      120      125
Ile Lys Gln Ile Arg Arg Leu Phe Val Asp Lys Phe Leu Glu Ser Tyr
      130      135      140
Pro Arg Glu Ser Lys Glu Ile Val Leu Asp Ile Asp Ala Trp Asp Ala
      145      150      155      160
Leu Thr His Gly His Gln Gln Leu Ser Leu Phe Asn Gly Tyr His Arg
      165      170      175
His Asp Ile Tyr Phe Pro Val Leu Ile Asn Glu Ala Ser Ser Gly Tyr
      180      185      190
Pro Leu Val Leu Gln Leu Arg Ala Gly Asn Ser His Ser Gly Lys Gly
      195      200      205
Val Ala Gly Ile Leu Lys Trp Leu Phe Leu Arg Ile Lys Arg Ala Leu
      210      215      220
Pro Glu Ile Arg Ile Val Leu Arg Gly Asp Gly Gly Phe Ser Leu Pro
      225      230      235      240
Glu Ile Ile Glu Val Cys Glu Lys Ser Gly Val Gly Tyr Val Phe Gly
      245      250      255
Phe Ser Asn Asn Asp Val Leu Lys Arg Lys Ile Asn Tyr Leu Leu Asp
      260      265      270
Arg Ala Arg Leu Glu Tyr Cys Arg Thr Gly Glu Lys Val Arg Leu Phe
      275      280      285
Asp Asp Val Tyr Tyr Ala Ala Arg Ser Trp Ser Glu Pro Arg Arg Val
      290      295      300
Ile Met Lys Ala Glu Trp Leu Glu Lys Gly Pro Asn Pro Arg Phe Ile
      305      310      315      320
Ile Thr Ser Leu Glu Thr Glu Ala Gln Asp Leu Tyr Asp Lys Phe Tyr
      325      330      335
Val Gln Arg Gly Ala Thr Ser Glu His Arg Ile Lys Glu Leu Lys Leu
      340      345      350
Gly Ile Lys Ser Asp Arg Leu Ser Cys Glu Lys Phe Ile Val Asn Gln
      355      360      365
Phe Arg Leu Phe Leu Ser Gln Ala Ala Tyr Ile Leu Met Leu Gly Ile
      370      375      380
Arg Gln Ala Ala Gln Gly Thr Lys Leu Ala Lys Ala Gln Val Pro Arg
      385      390      395      400
Leu Arg Glu Thr Ile Ile Lys Ile Ala Ala Lys Val Thr Val Ser Ala
      405      410      415
Arg Arg Val Leu Val Glu Leu Pro Tyr Tyr Cys Pro Phe Ser Ser Glu
      420      425      430
Ile Asn Leu Ile Met Glu Arg Leu Ala Ser Glu Phe Glu Ile Ile Phe
      435      440      445
Ser

```

- 5 <210> 60
 <211> 666
 <212> ADN
 <213> *Cylindrospermopsis raciborskii* T3

10 <400> 60

```

ctatctttgc cctgtaacaa tgtatgctac cctttgacca atattagtag catgatctgc 60
cattctctct aaacactgaa ttgctaattg taatagtaaa atgggctcca ctacccggg 120
aacatctttc tgctgcgcca aattacgata taactttttg taagcatcat ctactgtatc 180
atctaataat ttaatccttc taccactaat ctcgtctaaa tccgctaaag ctactaggct 240
ggtagccaac atagattggg catgatcgga cataatggca acctcccca aagtaggatg 300
ggggggatag ggaatatatt tcattgctat ttctgccaaa tctttggcat agtccccaat 360
acgttccaag tctctaacta attgcatgaa tgagcttaaa caccgagatt cttgggtctgt 420
gggagcttga ctgctcataa ttgtggcaca atcgacttct atttgtctgt agaagcgatc 480
aatttttttg tctaattctcc gtatttgctc agctgctgtt aaatcccgat tgaatagagc 540
ttggtgactc agacggaatg actgctctac taaagcacc atacgcaaaa catctcgttc 600
cagtctttta atggcacgta taggttgagg tttttcaaaa attgtatatt tcacaacagc 660
tttcat

```

<210> 61

<211> 221

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 61

```

Met Lys Ala Val Val Lys Tyr Thr Ile Phe Glu Lys Pro Gln Pro Ile
1      5      10      15
Arg Ala Ile Lys Arg Leu Glu Arg Asp Val Leu Arg Met Gly Ala Leu
      20      25      30
Val Glu Gln Ser Phe Arg Leu Ser His Gln Ala Leu Phe Asn Arg Asp
      35      40      45
Leu Thr Ala Ala Glu Gln Ile Arg Arg Leu Asp Lys Lys Ile Asp Arg
      50      55      60
Phe Tyr Arg Gln Ile Glu Val Asp Cys Ala Thr Ile Met Ser Ser Gln
65      70      75      80
Ala Pro Thr Asp Gln Glu Ser Arg Cys Leu Ser Ser Phe Met Gln Leu
      85      90      95
Val Arg Asp Leu Glu Arg Ile Gly Asp Tyr Ala Lys Asp Leu Ala Glu
      100     105     110
Ile Ala Met Lys Ile Phe Pro Tyr Pro Pro His Pro Thr Leu Gly Glu
      115     120     125
Val Ala Ile Met Ser Asp His Ala Gln Ser Met Leu Ala Thr Ser Leu
      130     135     140
Val Ala Leu Ala Asp Leu Asp Glu Ile Ser Gly Arg Arg Ile Lys Leu
145     150     155     160
Leu Asp Asp Thr Val Asp Asp Ala Tyr Lys Lys Leu Tyr Arg Asn Leu
      165     170     175
Ala Gln Gln Lys Asp Val Pro Gly Val Val Glu Pro Ile Leu Leu Leu
      180     185     190
Thr Leu Ala Ile Gln Cys Leu Glu Arg Met Ala Asp His Ala Thr Asn
      195     200     205
Ile Gly Gln Arg Val Ala Tyr Ile Val Thr Gly Gln Arg
10      210     215     220

```

<210> 62

<211> 1353

<212> ADN

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 62

```

tcagaaatat cgcgccatcat gttgaaccac ctggggaaga tgaatttgta tccaagcacc 60
accggtatca ggatgggttca tggccctgat ttgcccacca tgagctataa ttatttggcg 120
gacaatggat aaccctaaac cactaccagt aatttctact gtttcattct cagagcggga 180
ctcgcggtgt ctagctttgt cccccgata aaatctttga aagacatggg gtagatccat 240
gggagcaaat ccaaccccg aatcaataat gttaatttct aaaatctgat ttgatacttg 300
gtttaatat gtatctgctt ctggatcaac cccattaata gacttctccc cacaaactgg 360
attcatttca atgaaaatag taccgttcag gttgctgtat ttaatacagt tatctaacag 420
attaagaaac acttgataaa ttctggactt atcagcacat atatagacct tttccgggcc 480
ggagtaagaa atactaagat gctgattagc ggctaggggc tctaaattct cccagactga 540
aaaaattagg gagcggactt ctagcatttc caaattcagt tgtatggagg aggttatttc 600
catctgggtc aggtctaacc aattttggac taaattaatt agtctgtcaa cctcctgcat 660
caagcggatg acccaacggt ttagaggggg atctaagcga gtttgcaggg tttctgcgac 720
cagacgaatg gaagtcagag gtgttctcag ttcatgggcc aggtctgaaa aagagcggtc 780
acgttgctga tgaatgtcta caaattgttg gtgactttct agaacaacac ccacttgtcc 840
ccccggtagg ggaaaactgt tagctgctaa agacaatggc tttaatccta aaataccctg 900
accatgatct cgggaagggt gaaaaatcca ctcttgcat tgcggttttt gccaatcccg 960
ggtttgctca attaatgat ccagtcata ggatctcact aattccagta gcaggcgcac 1020
ttgaccgggt tgccatcttt gtaaatacag catttcccg cgcactgat tacaccatag 1080
tagttggttt tcttcatcta cttgtaaata tcccaaaggc gcagcatcca gcaactgttc 1140
ataagctttg agtgacaagc gtaagttttg ttgctcatct ctaacggtag atattttacg 1200
atgtaatcca gctaataagg gtaataatat cttttcagcg tgagggttta agggttgggt 1260
taactgctcc aaatgactgt taagttgaaa ttgttgcaa agccaaaaac caaaaccgac 1320
tgccaaaccc agaagaaatc ccaataagaa cat 1353

```

<210> 63

<211> 450

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* T3

<400> 63

```

Met Phe Leu Leu Gly Phe Leu Leu Gly Leu Ala Val Gly Phe Gly Phe
1          5          10          15
Trp Leu Trp Gln Gln Phe Gln Leu Asn Ser His Leu Glu Gln Leu Thr
20          25          30
Gln Pro Leu Asn Pro His Ala Glu Lys Ile Leu Leu Pro Leu Leu Ala
35          40          45
Gly Leu His Arg Lys Ile Ser Thr Val Arg Asp Glu Gln Gln Asn Leu
50          55          60
Arg Leu Ser Leu Lys Ala Tyr Glu Gln Leu Leu Asp Ala Ala Pro Leu
65          70          75          80
Gly Tyr Leu Gln Val Asp Glu Glu Asn Gln Leu Leu Trp Cys Asn Gln
85          90          95
Cys Ala Arg Glu Met Leu Tyr Leu Gln Arg Trp Gln Pro Gly Gln Val
100         105         110
Arg Leu Leu Leu Glu Leu Val Arg Ser Tyr Glu Leu Asp Gln Leu Ile
115         120         125
Glu Gln Thr Arg Asp Trp Gln Lys Pro Gln Met Gln Glu Trp Ile Phe
130         135         140
His Pro Ser Arg Asp His Gly Gln Gly Ile Leu Gly Leu Lys Pro Leu
145         150         155         160
Ser Leu Ala Ala Asn Ser Phe Pro Leu Pro Gly Gly Gln Val Gly Val
165         170         175
10 Phe Leu Glu Ser His Gln Gln Phe Val Asp Ile His Gln Gln Arg Asp

```

ES 2 525 045 T3

```

      180      185      190
Arg Ser Phe Ser Asp Leu Ala His Glu Leu Arg Thr Pro Leu Thr Ser
      195      200      205
Ile Arg Leu Val Ala Glu Thr Leu Gln Thr Arg Leu Asp Pro Pro Leu
      210      215      220
Asn Arg Trp Val Ile Arg Leu Met Gln Glu Val Asp Arg Leu Ile Asn
      225      230      235      240
Leu Val Gln Asn Trp Leu Asp Leu Thr Gln Met Glu Ile Thr Ser Ser
      245      250      255
Ile Gln Leu Asn Leu Glu Met Leu Glu Val Arg Ser Leu Ile Phe Ser
      260      265      270
Val Trp Glu Asn Leu Glu Pro Leu Ala Ala Asn Gln His Leu Ser Ile
      275      280      285
Ser Tyr Ser Gly Pro Glu Lys Val Tyr Ile Cys Ala Asp Lys Ser Arg
      290      295      300
Ile Tyr Gln Val Phe Leu Asn Leu Leu Asp Asn Cys Ile Lys Tyr Ser
      305      310      315      320
Asn Leu Asn Gly Thr Ile Phe Ile Glu Met Asn Pro Val Cys Gly Glu
      325      330      335
Lys Ser Ile Asn Gly Val Asp Pro Glu Ala Asp Thr Ile Leu Asn Gln
      340      345      350
Val Ser Asn Gln Ile Leu Glu Ile Asn Ile Ile Asp Ser Gly Val Gly
      355      360      365
Phe Ala Pro Met Asp Leu Pro His Val Phe Gln Arg Phe Tyr Arg Gly
      370      375      380
Asp Lys Ala Arg His Arg Glu Ser Arg Ser Glu Asn Glu Thr Val Glu
      385      390      395      400
Ile Thr Gly Ser Gly Leu Gly Leu Ser Ile Val Arg Gln Ile Ile Ile
      405      410      415
Ala His Gly Gly Lys Ile Arg Ala Met Asn His Pro Asp Thr Gly Gly
      420      425      430
Ala Trp Ile Gln Ile His Leu Pro Gln Val Val Gln His Asp Gly Gly
      435      440      445
Tyr Phe
      450

```

<210> 64

<211> 819

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 64

```

tcaaccaa ctatagccaa aaccctaac tgtgacaata tattctggat ggctagggtc 60
taactcta ttttcctca gccatcgaat gtgaacatcc accgttttac tgtcaccaac 120
aaaatcagga ccccaaacct ggtctaataa ctgttcccg gaccacacc tgcgagcata 180
actcataaat agttctagta accggaattc tttcggtgac aagctcacct ccctccctct 240
cactaacacc cgacattcct gaggatttaa actgatatcc ttatatatta aagtgggtat 300
caagggcaaa ttagaaaacc gctgacgacg taacagggcg cgacacctag ccaccatttc 360
ccgtacgcta aaaggcttag ttaggtaatc atccgccct acctctaaac ccagcaccg 420
gtcagtttca ctaccttcg cactcagaat taaaatcgg atggaattac cctgggtgacg 480
taacaaacga caaatatcta atccgttgat ttgtggcaac atcaagtcta gcacaagcag 540
gtcgaaggat aactcaccag gttgggtctc taaattcctg attaatcca cagcacaacg 600
accatcctta gcagtcacaa cttcataacc ttcacctct aaggctacta caagcatctc 660
tcggatcagt tcttcgtctt ccactattaa aacgcgacta actggttcaa tatccgattt 720
agtgaagtat ctagggtaat tcagtagtat acattgataa caaaaatttg taagaatgta 780
ctggtctggg tttccacta gtatatgat ctcactcat 819

```

10 <210> 65

<211> 272

<212> PRT

<213> *Cylindrospermopsis raciborskii* T3

15 <400> 65

ES 2 525 045 T3

```

Met Ser Glu Asp His Ile Leu Val Gly Asn Pro Asp Gln Tyr Ile Leu
1          5          10          15
Thr Asn Phe Cys Tyr Gln Cys Ile Leu Leu Asn Tyr Pro Arg Tyr Phe
          20          25          30
Thr Lys Ser Asp Ile Glu Pro Val Ser Arg Val Leu Ile Val Glu Asp
          35          40          45
Glu Glu Leu Ile Arg Glu Met Leu Val Val Ala Leu Glu Gly Glu Gly
          50          55          60
Tyr Glu Val Val Thr Ala Lys Asp Gly Arg Cys Ala Val Glu Leu Ile
65          70          75          80
Arg Asn Leu Glu Thr Gln Pro Gly Glu Leu Ser Phe Asp Leu Leu Val
          85          90          95
Leu Asp Leu Met Leu Pro Gln Ile Asn Gly Leu Asp Ile Cys Arg Leu
          100          105          110
Leu Arg His Gln Gly Asn Ser Ile Pro Ile Leu Ile Leu Ser Ala Lys
          115          120          125
Gly Ser Glu Thr Asp Arg Val Leu Gly Leu Glu Val Gly Ala Asp Asp
          130          135          140
Tyr Leu Thr Lys Pro Phe Ser Val Arg Glu Met Val Ala Arg Cys Arg
145          150          155          160
Ala Leu Leu Arg Arg Gln Arg Phe Ser Asn Leu Pro Leu Ile Pro Thr
          165          170          175
Leu Lys Tyr Lys Asp Ile Ser Leu Asn Pro Gln Glu Cys Arg Val Leu
          180          185          190
Val Arg Gly Arg Glu Val Ser Leu Ser Pro Lys Glu Phe Arg Leu Leu
          195          200          205
Glu Leu Phe Met Ser Tyr Ala Arg Arg Val Trp Ser Arg Glu Gln Leu
          210          215          220
Leu Asp Gln Val Trp Gly Pro Asp Phe Val Gly Asp Ser Lys Thr Val
225          230          235          240
Asp Val His Ile Arg Trp Leu Arg Glu Lys Leu Glu Leu Asp Pro Ser
          245          250          255
His Pro Glu Tyr Ile Val Thr Val Arg Gly Phe Gly Tyr Arg Phe Gly
          260          265          270

```

<210> 66

<211> 774

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* T3

<400> 66

```

tcaggcaaaa cgagagaagt ctaaagtggg tggaatatcc tgaattcttc caggacctat    60
agccccgtagt gcttctggta aactaatatc cccagtatat agggctttac ccacaattac    120
tcctgtaacc cctgatgtt ctaaagataa taagggttaat aggtcagtaa cagaaccac    180
acccccagag gcaatcacgg gtatggaaat agcagatacc aagtctctta atgctcgcaa    240
gtttggtccc tgaagcgtag catcacggtt tatatccgta taaataatag ctgccgcacc    300
caattcctgc atttgggttg ctagttgggg ggccaaaatt tgagaagttt ctaaccaacc    360
cctggtagca actagaccat tccgcgcatac aatcccaatt ataatttgct gggggaattg    420
ttcacacagt ccttgaacca gatctgggtg ctctactgct acagttccca gaattgccca    480
ctgtacccca agattaaata actgtataac gctggagcta tcacgtattc ctccgccaac    540
ttcaataggt atggaaatag cattggtaat agcttctata gtagataaat taactatttt    600
accagttttt gctccatcta aatctactaa atgtagtctt gttgctcctt ggtctgcca    660
cattttagcg gtttcacag ggttatggct gtaaacctgg gattgtgcat agtcaccttt    720
gtagagtctt acacaacgcc cctctaatag atctattgct gggataactt ccat          774

```

<210> 67

<211> 257

<212> PRT

15 <213> *Cylindrospermopsis raciborskii* T3

<400> 67

ES 2 525 045 T3

Met Glu Val Ile Pro Ala Ile Asp Leu Leu Glu Gly Arg Cys Val Arg
1 5 10 15
Leu Tyr Lys Gly Asp Tyr Ala Gln Ser Gln Val Tyr Ser His Asn Pro
20 25 30
Val Glu Thr Ala Lys Met Trp Ala Asp Gln Gly Ala Thr Arg Leu His
35 40 45
Leu Val Asp Leu Asp Gly Ala Lys Thr Gly Lys Ile Val Asn Leu Ser
50 55 60
Thr Ile Glu Ala Ile Thr Asn Ala Ile Ser Ile Pro Ile Glu Val Gly
65 70 75 80
Gly Gly Ile Arg Asp Ser Ser Ser Val Ile Gln Leu Phe Asn Leu Gly
85 90 95
Val Gln Trp Ala Ile Leu Gly Thr Val Ala Val Glu Gln Pro Asp Leu
100 105 110
Val Gln Gly Leu Cys Glu Gln Phe Pro Gln Gln Ile Ile Ile Gly Ile
115 120 125
Asp Ala Arg Asn Gly Leu Val Ala Thr Arg Gly Trp Leu Glu Thr Ser
130 135 140
Gln Ile Leu Ala Pro Gln Leu Ala Thr Gln Met Gln Glu Leu Gly Ala
145 150 155 160
Ala Ala Ile Ile Tyr Thr Asp Ile Asn Arg Asp Gly Thr Leu Gln Gly
165 170 175
Pro Asn Leu Arg Ala Leu Arg Asp Leu Val Ser Ala Ile Ser Ile Pro
180 185 190
Val Ile Ala Ser Gly Gly Val Gly Ser Val Thr Asp Leu Leu Thr Leu
195 200 205
Leu Ser Leu Glu His Gln Gly Val Thr Gly Val Ile Val Gly Lys Ala
210 215 220
Leu Tyr Thr Gly Asp Ile Ser Leu Pro Glu Ala Leu Arg Ala Ile Gly
225 230 235 240
Pro Gly Arg Ile Gln Asp Ile Pro Pro Thr Leu Asp Phe Ser Arg Phe
245 250 255
Ala

<210> 68

<211> 396

5 <212> ADN

<213> Cylindrospermopsis raciborskii T3

<400> 68

atgagttggg ccacaatgaa ggacgtcttg attttaatag tcaaatccct ccaaattccat 60
tataatccca tgaatgctct ttcaattcct acctggatta tccatatttc tagtggtcatt 120
gaatgggtag ttgccatttc cctcatctgg aaatatggcg aactgaccca aaaccatagt 180
tggaggggat ttgccttagg tatgataccc gccttaatta gcgccctatc cgcttgtacc 240
tggcattatt tcgataatcc ccagtcoccta gaatgggttag tcaccctcca ggctactact 300
acgttaatag gtaattttac tctttgggca gcagcagtcct gggtttggcg ttctactcga 360
10 ccgaatgagg ttctcagtat ctcaaataag gagtag 396

<210> 69

<211> 131

<212> PRT

15 <213> Cylindrospermopsis raciborskii T3

ES 2 525 045 T3

<400> 69
Met Ser Trp Ser Thr Met Lys Asp Val Leu Ile Leu Ile Val Lys Ser
1 5 10 15
Leu Gln Ile His Tyr Asn Pro Met Asn Ala Leu Ser Ile Pro Thr Trp
20 25 30
Ile Ile His Ile Ser Ser Val Ile Glu Trp Val Val Ala Ile Ser Leu
35 40 45
Ile Trp Lys Tyr Gly Glu Leu Thr Gln Asn His Ser Trp Arg Gly Phe
50 55 60
Ala Leu Gly Met Ile Pro Ala Leu Ile Ser Ala Leu Ser Ala Cys Thr
65 70 75 80
Trp His Tyr Phe Asp Asn Pro Gln Ser Leu Glu Trp Leu Val Thr Leu
85 90 95
Gln Ala Thr Thr Thr Leu Ile Gly Asn Phe Thr Leu Trp Ala Ala Ala
100 105 110
Val Trp Val Trp Arg Ser Thr Arg Pro Asn Glu Val Leu Ser Ile Ser
115 120 125
Asn Lys Glu
130

<210> 70
<211> 20
5 <212> ADN
<213> Artificial

<220>
10 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 70

ttaattgctt ggtctatctc 20

15 <210> 71
<211> 20
<212> ADN
<213> Artificial

20 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 71

25 caataccgaa gaggagatag 20

<210> 72
<211> 20
<212> ADN
30 <213> Artificial

<220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3

35 <400> 72

taggcgtgtt agtgggagat 20

<210> 73
40 <211> 20
<212> ADN
<213> Artificial

<220>
45 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 73

tgtgtaacca atttgtgagt 20

<210> 74
 <211> 20
 <212> ADN
 5 <213> Artificial

 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 10 <400> 74
 ttagccggat tacagtgaa 20

 <210> 75
 15 <211> 20
 <212> ADN
 <213> Artificial

 <220>
 20 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3

 <400> 75
 ctggactcgg cttgttgctt 20
 25
 <210> 76
 <211> 20
 <212> ADN
 <213> Artificial
 30
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3

 <400> 76
 35 cagcgagtta caccaccac 20

 <210> 77
 <211> 20
 40 <212> ADN
 <213> Artificial

 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 45
 <400> 77
 ctgcactaa atattctacc 20
 50
 <210> 78
 <211> 19
 <212> ADN
 <213> Artificial
 55
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3

 <400> 78
 60 aaaacctcag ctccacaa 19

 <210> 79
 <211> 22
 <212> ADN
 65 <213> Artificial

<220>

<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3

<400> 79

5

atgattttgg aggtccattg tt

22

<210> 80

<211> 42156

10

<212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 80

gtttttactg	caaaagcata	ttcatattat	attctaatag	ggttggtgga	atattcaagg	60
ggaggttaga	aaatgcatc	gctcttatga	atgaggttgt	ctatccgaat	atcaaataatt	120
ggtggttgaa	aaaagacctt	atatgctggac	acagattccc	atgatgaaaa	tatatcattg	180
tcaagtcaat	tagtcaaccc	cccaatagac	atctccgaaa	aagaatcaaa	gtgtgataaa	240
atttgcagta	cagcaggata	taaaatagtt	tttctcttat	acttctgagt	gtaggcttgc	300
gtccgcctcc	ggcgccacgt	ttgcggtttg	ctaaggagt	aaacacggtg	cgtaaatatg	360
tatcagcaac	ctgagataac	agctcgttga	atgcttagcg	gttaagtcca	gtcattgctc	420
gtagcagtcg	ctcttgattc	aggatgcggt	ctaagttcaa	cattaatgtc	accctacttg	480
tctgcttgat	tattatccct	tattttccaa	caactcta	gaaagtacct	ataacagcaa	540
acgaagatgc	agctacatta	cttcagcgtg	ttggactgtc	cctaaaggaa	gcacaccaac	600
aacttgaggc	aatgcaacgc	cgagcgcacg	aaccgatcgc	aattgtgggg	ctggggctgc	660
ggtttccggg	agctgattca	ccacagacat	tctggaaact	acttcagaat	ggtgttgata	720
tggtcaccga	aatccctagc	gatcgtctgg	cagttgatga	atactatgat	ccccaacctg	780
ggtgtccagg	caaaatgtat	attcgtgaag	ccgcttttgt	tgatgcagtg	gataaattcg	840
atgcctcggt	ttttgatatt	tcgccacgtg	aagcggccaa	tatagatccc	cagcatagaa	900
tggtgctgga	ggtagccttg	gaggcactcg	aaagggctgg	cattgctccc	agccaattga	960
tggatagcca	aacgggggta	tttgcgggga	tgagcgaaaa	tgactattat	gctcacctag	1020
aaaatacagg	ggatcatcat	aatgtctatg	cggcaacggg	caatagcaat	tactatgtct	1080
cggggcggtt	atcctatcta	ttggggcttc	aaggacctaa	catggtcggt	gatagtgcct	1140
gttctcctc	cttagtggct	gtacatcttg	cctgtaatag	tttgcggatg	ggagaatgtg	1200

atctggcact	ggctgggtgg	gttcagctta	tgtaaatccc	agaccctatg	attgggactg	1260
cccagttaaa	tgctcttgcg	accgatggtc	gtagtaaaac	atctgacgct	gccgccgatg	1320
gctatggacg	cggcgaaagt	tgtggcatga	ttgtacttaa	aagaataagt	gacgcgatcg	1380
tggcagacga	tccaatttta	gccgtaatcc	ggggtagtgc	agtcacatcat	ggcgggctga	1440
gcagtggttt	aactgcccct	aataagctgt	ctcaagaagc	cttactgcgt	caggcactac	1500
aaaacgcaa	ggttcagccg	gaagcagtc	gttatatcga	agcccatggc	acagggacac	1560
aactgggcca	cccgatggag	gtgggagcat	taacgaccgt	ctttggatct	tctcgttcag	1620
aaccttgtg	gattggctct	gtcaaaacta	atctcgga	cctagaacca	gccgctggta	1680
ttgcggggtt	aataaaaagtc	attttatcat	tacaagaaaa	acagattcct	cccagttctc	1740
attttcaaaa	ccctaattccc	ttcattgatt	gggaatcttc	gccagttcaa	gtgccgacac	1800
agtgtgtacc	ctggactggg	aaagagcgcg	tcgctggagt	tagctcggtt	ggtatgagcg	1860
gtacaaactg	tcatctagtt	gtcgcagaag	cacctgtccg	ccaaaacgaa	aaatctgaaa	1920
atgcaccgga	gcgtccttgt	cacattctga	ccctttcagc	caaaaaccgaa	gcggcactca	1980
acgcattggt	agcccgttac	atggcatttc	tcaggggaagc	gcccgccata	tccctagctg	2040
atctttgtta	tagtgccaat	gtcgggctga	atctttttgc	ccatcgctta	agttttatct	2100
ccgagaacat	cgcgagttta	tcagaacaat	tagaacactg	cccacagcag	gctacaatgc	2160
caacgcaaca	taatgtgata	ctagataatc	aactcagccc	tcaaatcgct	tttctgttta	2220
ctggacaagg	ttcgagctac	atcaacatgg	ggcgtgagct	ttacgaaact	cagcccacct	2280
tccgtcggat	tatggacgaa	tgtgacgaca	ttctgcatcc	attgttgggt	gaatcaattc	2340
tgaacatact	ctacacttcc	cctagcaaac	ttaatcaaac	cgtttatacc	caacctgccc	2400
tttttgcttt	tgaatatgcc	ctagcaaaac	tatggatatc	atggggtatt	gagcctgatg	2460
tcgtactggg	tcacagcgtg	ggtgaatatg	tagccgcttg	tctggcgggt	gtcttttagt	2520
tagaagatgg	gttaaaactc	attgcatctc	gtggatgttt	gatgcaagcc	ttaccgccgg	2580
ggaaaatgct	tagtatcaga	agcaatgaga	tcggagtga	agcgtctatc	gcgccttata	2640
gtgcagaagt	atcaattgca	gcaatcaatg	gacagcaaa	cgtggtgatc	tccggcaaa	2700
ctgaaattat	agataattta	gcagcagagt	ttgcatcgga	aggcatcaaa	acacacctaa	2760
ttacagtctc	ccacgctttc	cactcgccaa	tcagtacccc	catgctgaaa	gcattccgag	2820
acgttgccag	caccatcagc	tataggctac	ccagtttatc	actgatttct	aacggtacag	2880
ggcaattggc	aacaaaggag	gttgctacac	ctgattattg	ggtgctgcat	gtccatttcta	2940
ccgtccggtt	tgccgatggt	attgccacat	tggcagaaca	gaatactgac	atcctcctag	3000
aagtaggacc	caaaccaata	ttgttgggta	tggcaaaagca	gatttatagt	gaaaacggtt	3060
cagctagtca	tccgctcttc	ctaccagtt	tgcgtaaga	tggcaacgat	tggcagcaga	3120
tgctttctac	ttgtggacaa	cttgtagtta	atggagtcaa	gattgactgg	gcgggttttg	3180
acaaggatta	ttcacgacac	aaaatattgt	tgccaccta	tccgtttcag	agagaacgat	3240
attggattga	aagctccgtc	aaaaagcccc	aaaaacagga	gctgcgcca	atgttgata	3300

15

agatgatccg	gctaccatca	gagaacaaag	tggtgtttga	aaccgagttt	ggcgtgcgac	3360
agatgcctca	tatctccgat	catcagatat	acggtgaagt	cattgtaccg	ggggcagtat	3420
tagcttcctt	aatcttcaat	gcagcgcagg	ttttataccc	agactatcag	catgaattaa	3480
ctgatattgc	tttttatcag	ccaattatct	ttcatgacga	cgatacgggtg	atcgtgcagg	3540
cgattttcag	ccctgataag	tcacaggaga	atcaaagcca	tcaaacattt	ccacccatga	3600
gcttccagat	tattagcttc	atgccggatg	gtcccttaga	gaacaaaccg	aaagtccatg	3660
tcacaggggtg	tctgagaatg	ttgcgcgatg	cccaaccgcc	aacactctcc	ccgaccgaaa	3720
tacgtcagcg	ctgtccacat	accgtaaatg	gtcatgactg	gtacaatagc	ttagtcaaac	3780
aaaaatttga	aatgggtcct	tccttttaggt	gggtacagca	actttggcat	ggggaaaatg	3840
aagcattgac	ccgtcttcac	ataccagatg	tggtcggctc	tgtatcagga	catcaacttc	3900
acggcatatt	gctcgatggg	tcactttcaa	ccaccgctgt	catggagtac	gagtacggag	3960
actccgcgac	cagagttcct	ttgtcatttg	cttctctgca	actgtacaaa	cccgtcacgg	4020
gaacagagtg	gtggtgctac	gcgaggaaga	ttggggaatt	caaatatgac	ttccagatta	4080
tgaatgaaat	cggggaaacc	ttgggtgaaag	caattggcct	tgtacttcgt	gaagcctctc	4140
ccgaaaaatt	cctcagaaca	acatacgtac	acaactggct	tgtagacatt	gaatggcaag	4200
ctcaatcaac	ttccctagtc	ccttctgatg	gcactatctc	tggcagttgt	ttggttttat	4260
cagatcagca	tggaaacagg	gctgcattgg	cacaaaggct	agacaatgct	ggagtgccag	4320
tgaccatgat	ctatgctgat	ctgatactgg	acaattacga	attaatatct	cgtactttgc	4380
cagatttaca	acaagtcgtc	tatttatggg	ggttggatca	aaaagaggat	tgtcacccca	4440
tgaagcaagc	agaggataac	tgtacatcgg	tgctatatct	tgtgcaagca	ttactcaata	4500
cctactcaac	cccgccatcc	ctgcttattg	tcacctgtga	tgcacaagcg	gtggttgaac	4560
aagatcgagt	aaatggcttc	gcccaatcgt	ctttgttggg	acttgccaaa	gttatcatgc	4620
tagaacaccc	agaattgtcc	tgtgtttaca	tggatgtgga	agccggatat	ttacagcaag	4680
atgtggcgaa	cacgatattt	acacagctaa	aaagaggcca	tctatcaaag	gacggagaag	4740
agagtacgtt	ggcttggcgc	aatggacaag	catactagc	acgtcttagt	caatataaac	4800
ccaaatccga	acaactgggt	gagatccgca	gcgatcgcag	ctatttgatc	actggtggac	4860
ggggcggtgt	cggcttacia	atcgcacggt	ggttagtgga	aaagggggct	aaacatctcg	4920
ttttgttggg	gcgcagtcag	accagttccg	aagtcagctc	ggtgttggat	gagctagaat	4980
cagccggggc	gcaaatcatt	gtggctcaag	ctgatattag	cgatgagaag	gtattagcgc	5040
agattctgac	caatctaacc	gtacctctgt	gtggtgtaat	ccacgccgca	ggagtgtttg	5100
atgatgcgag	tctactccaa	caaactccag	ccaagctcaa	aaaagttcta	ttgccaaaag	5160
cagagggggc	ttggattctg	cataatttga	ccctggagca	gcgactagac	ttctttgttc	5220
tcttttcttc	tgccagttct	ctattaggtg	cgccagggca	ggccaactat	tcagcagcca	5280
atgctttcct	agatggttta	gctgcctatc	ggcgagggcg	aggactcccc	tgtttgtcta	5340
tctgctgggg	ggcatgggat	caagtcgcta	tggctgcacg	acaagggcta	ctggacaagt	5400
taccgcaaa	agggtgaagag	gccatcccgt	tacagaaagg	cttagacctc	ttcggcgaat	5460
tactgaacga	gccagccgct	caaattggtg	tgatcccaat	tcaatggact	cgcttcttgg	5520
atcatcaaaa	aggtaatttg	cctttttatg	agaagtttct	taagtctagc	cggaaagcgc	5580
agagttacga	ttcgtatggc	gtcagtcaca	cagaagatat	tcagaggaaa	ctgaagcaag	5640
ctgctgtgca	agatcgacca	aaattattag	aagtgcatact	tcgctctcaa	gtcgctcaac	5700
tgtaggaat	aaactgtggc	gagctaccaa	atgaagaagg	aattggtttt	gttacattag	5760
gtcttgactc	gctcacctct	attgaactgc	gtaacagttt	acaacgcaca	ttagattgtt	5820
cattacctgt	cacctttgct	tttgactacc	caactataga	aatagcgggt	aagtacctaa	5880
cacaagttgt	aattgcaccg	atggaaagca	cagcatcgca	gcaaacagac	tctttatcag	5940
caatgttcac	agatacttcg	tccatcgga	gaattcttga	caacgaaaca	gatgtgttag	6000
acagcgaaat	gcaaagtgat	gaagatgaat	ctttgtctac	acttatacaa	aaattatcaa	6060
cacatttgga	ttaggagtga	tcaataatta	tacattgcgg	acgtgagcat	acaagtaaa	6120
gaaaaatgaa	tgaacgcttt	gtcagaaaat	caggtaactt	ctatagtcaa	gaaggcattg	6180
aacaaaaatg	aggagttaca	agccgaactt	gaccgtttaa	aatacgcgca	acgggaacca	6240
atcgccatca	ttggaatggg	ctgtcgcttt	cctggtgcag	acacacctga	agcttttttg	6300
aaattattgc	acaatggggt	tgatgtctatc	caagagattc	caaaaagccg	ttgggatatt	6360
gacgactatt	atgatccac	accagcaaca	cccggcaaaa	tgtatacacg	ttttggtggt	6420
tttctcgacc	aaatagcagc	cttcgaccct	gagttcttct	gcatttctac	tcgtgaggca	6480
atcagcttag	accctcaaca	gagattgctt	ctggaagtga	gttgggaagc	cttagaacgg	6540
gctgggctga	caggcaataa	actgactaca	caaacagggtg	tctttgttgg	catcagtga	6600
agtgattatc	gtgatttgat	tatgcgtaat	ggttctgacc	tagatgtata	ttctggttca	6660
ggtaactgcc	atagtacagc	cagcgggcgt	ttatcttatt	atttgggact	tactggaccc	6720

aatttgtccc	ttgataccgc	ctgttcgtcc	tctttggttt	gtgtggcatt	ggctgtcaag	6780
agcctacgtc	aacaggagtg	tgatttggca	ttggcgggtg	gtgtacagat	acaagtgata	6840
ccagatggct	ttatcaaagc	ctgtcaatcc	cgtatgttgt	cgcctgatgg	acgggtgcaaa	6900
acatttgatt	tccaggcaga	tggttatgcc	cgtgctgagg	gggtgtggat	ggtagttctc	6960
aaacgcctat	ccgatgcaat	tgctgacaat	gataatatcc	tgcccttgat	tcgtgggtgcc	7020
gcagtcaatc	atgatggcta	cacgagtgga	ttaacctgtc	ccagtgggtc	ctcacaacgg	7080
gcggtgatcc	aacaggcatt	agcggatgct	ggaatacacc	cggatcaaat	tagctatat	7140
gaggcacatg	gcacaggtag	atccttaggc	gatcctattg	aaatgggtgc	gattggggcaa	7200
gtctttggtc	aacgctcaca	gatgcttttc	gtcggttcgg	tcaagacgaa	tattgggtcat	7260
actgaggctg	ctgctgggtat	tgctgggtctc	atcaagggtg	tactctcaat	gcagcacagg	7320
gaaatcccag	caaacttaca	cttcgaccag	ccaagtcctt	atattaactg	ggatcaatta	7380
ccagtacgta	tcccaacaga	aacaatacct	tggtctacta	gcgatcgctt	tcgaggagtc	7440
agtagctttg	gcttttagtg	cacaaactct	catatcgtag	tagaggcagc	cccaaacata	7500
gagcaaccta	ctgatgatat	taatcaaagc	ccgcataatt	tgaccttagc	tgcaaaaaaca	7560
cccgcagccc	tgtaagaact	ggctcggcgt	tatgcgactc	agatagagac	ctctcccgat	7620
gttctctctg	cggacatttg	tttcacagca	cacatagggc	gtaaacattt	taaacatagg	7680
tttgcggtag	tcacgggaatc	taaagagcaa	ctgcgtttgc	aattggatgc	atttgacaaa	7740
tcaggggggtg	tggggcgaga	agtcaaatcg	ctaccaaaga	tagcctttct	ttttacaggt	7800
caaggctcac	agtatgtggg	aatgggtcgt	caactttacg	aaaaccaacc	taccttcga	7860
aaagcactcg	cccattgtga	tgacatcttg	cgtgctgggtg	catatttcga	ccgatcacta	7920
ctttcgattc	tctaccaga	gggaaatca	gaagccattc	accaaaccgc	ttatactcag	7980
cccgcgcttt	ttgctcttga	gtatgcgtag	gctcagttgt	ggcactcctg	gggtatcaaa	8040
ccagatatcg	tgatggggca	tagtgtaggt	gaatacgtcg	ccgcttgtgt	ggcgggcata	8100
ttttcttttag	aggatgggct	gaaactaatt	gctactcgtg	gtcgtctgat	gcaatcccta	8160
cctcaagacg	gaacgatggg	ttcttctttg	gcaagtgaag	ctcgtatcca	ggaagctatt	8220
acaccttacc	gagatgatgt	gtcaatcgca	gcgataaatg	ggacagaaaag	cgtgggttatc	8280
tctggcaaac	gcacctctgt	gatggcaatt	gctgaacaac	tcgccaccgt	tggtcatcaag	8340
acacgccaac	tgacgggttc	ccatgccttc	cattcaccac	ttatgacacc	catcttggat	8400
gagttccgcc	agggtggcag	cagtatcacc	tatcaccagc	ccaagttgct	acttgtctcc	8460
aacgtctccg	ggaaagtggc	cggccctgaa	atcaccagac	cagattactg	ggtacgccat	8520
gtccgtgagg	cagtgcgctt	tgccgatgga	gtgaggacgc	tgaatgaaca	agggtgtcaat	8580
atctttcttg	aaatcgggtc	taccgctacc	ctggttggga	tggcactgcg	agtaaatgag	8640
gaagattcaa	atgcctcaaa	aggaaactcg	tcttgctacc	tgcccagttt	acgggaaagc	8700
cagaaggatt	gtcagcagat	gttactagt	ctgggtgagt	tgtacgtaca	tggatatgat	8760
attgatggg	gtgcatttaa	tcggggatat	caaggacgca	agggtgatatt	gccaacctat	8820
ccgtttcagc	gacaacgtta	ttggcttccc	gaccctaagt	tggcacaag	ttccgattta	8880
gatacctttc	aagctcagag	cagcgcatca	tcacaaaatc	ctagcgtgtg	gtccacttta	8940
ctgatggaa	atltgcaagc	agggtatgtc	caatctttag	ttgggctttt	ggatgatgaa	9000
cggaaactct	ctgctgctga	acgaattgca	ctaccagta	ttttggagtt	tttggtagag	9060
gaacaacagc	gacaaataag	ctcaaccaca	actcctcaaa	cagtttttaca	aaaaataagt	9120
caaacttccc	atgaggacag	atatgaaata	ttgaagaacc	tgatcaaatc	tgaaatcgaa	9180
acgattatca	aaagtgttcc	ctccgatgaa	caaatgtttt	ctgacttagg	aattgattcc	9240
ttgatggcga	tcgaactgcg	taataagctc	cgttctgcta	tagggttgga	actgccagtg	9300
gcaatagtat	ttgaccatcc	cacgattaag	cagttaacta	acttcgtact	ggacagaatt	9360
gtgccgcagg	cagaccaaaa	ggacgttccc	accgaatcct	tgtttgcttc	taaacaggag	9420
atatcagttg	aggagcagtc	ttttgcaatt	accaagctgg	gcttatcccc	tgttcccac	9480
tccctgcata	ttcctccatg	gacggttaga	cctgcggtaa	tgccagatgt	aacaaaacta	9540
agccaacttg	aaagagaggg	ctatggctgg	atcggaag	gagcgatcgc	cccgcacct	9600
ctcattgccg	atcgcatcaa	tttactcaac	agtgggtgata	tgcttgggtt	ctgggtaatg	9660
gagcgatcag	gagagtggg	cgcgtggcag	gtgctacaac	cgacatctgt	tgatccatat	9720
acttatggaa	gttgggatga	agtaactgac	caaggtaaac	tgcaagcaac	cttcgacca	9780
agtggacgca	atgtgtatat	tgctcggggt	gggtctagca	acctccccac	ggtagccagc	9840
cacctcatga	cgttcagac	tttattgatg	ctgcccggaa	ctggtcgtga	cacaatcttt	9900
gtctgtctgg	caatgccagg	ttatgccaaa	taccacagtc	aaacaggaaa	atcgccggaa	9960
gagtatattg	cgtcgactga	cgaggatggg	atcccaatgg	acgagtttat	tgcactttct	10020
gtctacgact	ggcctgttac	cccatcggtt	cgtgttctgc	gagacgggta	tccacctgat	10080
cgagattctg	gtgggtcacgc	agttagtacg	gttttccagc	tcaatgattt	cgatggagcg	10140

atcgaagaaa	catatcgtcg	tattatccgc	catgccgatg	tccttgggtct	cgaaagagggc	10200
taaatttcag	gcgttgggtga	atagaaccca	cattccgcag	ataaggtctt	atgaataaaa	10260
aacaggtaga	cacattgtta	atacacgctc	atctttttac	catgcagggc	aatggccttg	10320
gatatattgc	cgatggggca	attgcggttc	agggtagcca	gatcgtagca	gtggattcga	10380
cagaggcctt	gctgagtcac	tttgaaggaa	ataaaacaat	taatgcggta	aattgtgcag	10440
tgttgccctg	actaattgat	gctcatatac	atacgacttg	tgctattctg	cgtggagtgg	10500
cacaggatgt	aaccaattgg	ctaattggacg	cgacaattcc	ttatgcactt	cagatgacac	10560
ccgcagtaaa	tatagccgga	acgcgcttga	gtgtactcga	agggctgaaa	gcaggaacaa	10620
ccacattcgg	cgattctgag	actccttacc	cgctctgggg	agagtttttc	gatgaaattg	10680
gggtacgtgc	tattctatcc	cctgccttta	acgcctttcc	actagaatgg	tcggcatgga	10740
aggagggaga	cctctatccc	ttcgatatga	aggcaggacg	acgtggtatg	gaagaggctg	10800
tggattttgc	ttgtgcatgg	aatggagccg	cagagggacg	tatcaccact	atgttgggac	10860
tacaggcggc	ggatatgcta	ccactggaga	tcctacacgc	agctaaagag	attgccccac	10920
gggaaggcct	aatgctgcat	attcatgtgg	cccaggggaga	tcgagaaaca	aaacaaattg	10980
tcaaacgata	tggttaagcgt	ccgatcgcac	ttctagctga	aattggctac	ttggacgaac	11040
agtgtgctgg	agttcacctc	accgatgccca	cagatgaaga	agtgatacaa	gtagccaaaa	11100
gtggtgctgg	catggcactc	tgctcgggcg	ctattggcat	cattgacggt	ctgtgtccgc	11160
ccgctcatgt	ttttcgacaa	gcaggcgggt	ccgttgactc	cggttctgat	caagcctgtg	11220
gcaacaactg	ttgtaacatc	ttcaatgaaa	tgaagctgac	cgccttattc	aaacaaataa	11280
aatatcatga	tccaaccatt	atgccggcct	gggaagtcct	gcgtatggct	accatcgaag	11340
gagcgagggc	gatttggttta	gatcacaaaga	ttggctctct	tcaagtgggc	aaagaagccg	11400
acctgatctt	aatagacctc	agttccccta	acctctcgcc	caccctgtct	aacctatttc	11460
gtaaccttgt	acctaacttg	gtgtatgctg	cttcaggaca	tgaagttaa	agcgtcatgg	11520
tggcgggaaa	acttttagtg	gaagactacc	aagtcctcac	ggtagatgag	tccgctattc	11580
tcgctgaagc	gcaagtacaa	gctcaacaac	cttgccaacg	tgtgaccgct	gacccatttc	11640
acaaaaagat	gggtgttaag	gaagcgatgg	cctaagggtaa	attatagata	caggcttatc	11700
tgcaacaaca	tttctgaatc	aaacctggag	gggcaaacca	atgaccatat	atgaaaataa	11760
gttgagtagt	tatcaaaaaa	atcaagatgc	cataatatct	gcaaaagaac	tcgaagaatg	11820
gcattttaatt	ggacttctag	accattcaat	agatgcggta	atagtagcca	attattttct	11880
ttagcaagag	tgtatgacaa	tttcagagag	aataaaaaag	agtaaatatt	ttagcgctta	11940
tcccggtcat	ccatcagtaa	gtagcttggg	acaagagtgt	tatgaatgag	aaagttagct	12000
tgaattagca	aagtatcaag	aagacgcacc	cacattgatt	aaagaaatgc	ggaggctggt	12060
acatccgtac	ataagtccaa	ttgatagact	taggggtgaa	gttgatgata	tttggagtta	12120
tggctgtaat	ttagcaaaac	ttggtgataa	aaaactgttt	gcgggtatcg	ttagagagtt	12180
taaagaagat	aaccttgccg	caccacattg	tgacgtaatg	gcattggggt	ttctcgaata	12240
ttataaagat	aaaccaataa	tcataaatca	aatcgcaaga	aatgtatatt	taaaaacgtc	12300
tgcatcagga	ggagaaatag	tgctttggga	tgaatggcca	actcaaagcg	aatatatagc	12360
atacaaaaaca	gatgatccag	ctagtttcgg	tcttgatagc	aaaaagatcg	cacaaccaa	12420
acttgagatc	caaccgaacc	agggagattt	aattctattc	aattccatga	gaattcatgc	12480
ggtgaaaaag	atagaaactg	gtgtacgtat	gacatgggga	tggttgattg	gatactctgg	12540
aactgataaa	ccgcttggtta	tttggactta	atgtagcggt	tccatttgag	tcaaggcacg	12600
agaagcttct	aaagctggaa	tagatacact	atcattctca	actacactct	caaagtctct	12660
aggtaactgt	gccccaaaca	tcagcattcc	aatggcggtg	aacaaaaaga	aagccaacca	12720
caagatatgg	ttactctcaa	atttaacagc	agctacatcc	gcaggtaaaa	atcctacacc	12780
aaacgcgatt	aagttaacat	tgccggagagt	atgcccttga	gccaaaccca	agaagtaccc	12840
acatagtatg	caacatactg	aattgcatac	taggacaagt	accaaccagg	gaataaaaaa	12900
atcaatatct	tcaataattt	ctgcgtgggt	gggttaacaac	ccaaaaacat	catcgggaaa	12960
tagccaacac	gctccgccga	aaaccagact	cactagcaga	gccattccca	cagaaacttt	13020
tgccagaggt	gctaactggt	ctgtggctcc	tttcccttta	aaatttcctg	ccagagtttc	13080
tgtacagaat	cccaatcctt	caacaatgta	gatgctcaaa	gccccatatc	gtaagagcaa	13140
ggcattttga	gcgtagataa	ttgtcccat	ttgtgcccct	tcgtagttaa	acgttaagtt	13200
ggtaaacata	caaactaaat	tgctgacaaa	gatgtttcca	ttgagagtta	aggtggagcg	13260
tatagctttt	atgtcccaaa	tttttccagc	taattctttt	acctcttgcc	acgggatttc	13320
tttgagacaa	aaaaacaatc	ccaccaatag	ggtgagatat	tgacttgtag	cagaagctac	13380
tctgcccccc	atgctcgacc	agtctaagtg	gataataaac	aagtagtcga	gtgcgatatt	13440
ggcagcattg	cccacaaccg	acaacaacac	aactaagcca	tttttttccc	gtcccagaaa	13500
ccagccaagc	aggacaaagt	tgagcaaaat	ggcaggcgct	ccccaaactc	gggtgttaaa	13560

atacgcttga	gctgaagact	tcacctctgg	gccgacatct	agtatagaaa	accccaacac	13620
ccctaacggg	tactgtaaca	gtatgatcgc	cacccccagc	accagagcaa	ttaaaccatt	13680
aagcagtcgc	gccaacagta	cgccctctcg	gtcatctcgt	ccgactgctt	gtgctgttaa	13740
cgcagtggtg	cccattcgtg	aaaacgataa	aacaaagtag	agaaagttaa	gcagggtttcc	13800
agcaagggct	actccagcta	ggtagtggat	ttccgagaga	tgacctaaaga	acatgatact	13860
gactaaatta	ctcagtggtg	ctataatatt	cgataggacg	ttggtaaaaag	ctagtcggaa	13920
gtagcggggg	ataaagtcac	actggcttgg	aaatgtcagg	ctcataagat	taatttgaca	13980
gtagagtgtg	tggaaaataa	gggataataa	tcaagcagac	aagtaggggtg	acattaatgt	14040
tgaacttaga	ccgcatctcg	aatcaagagc	gactgctacg	agaaatgact	ggacttaacc	14100
gccaagcatt	caacgagctg	ttatctcagt	ttgctgatac	ctatgaacgc	accgtgttca	14160
actccttagc	aaaccgcaaa	cgtgcgcccg	ggggcgagcg	caagcctaca	ctcagaagta	14220
tagaggaaaa	actattttat	atcctgctgt	actgcaaagt	ttatccgacg	tttgacttgc	14280
tgagtgtgtt	gttcaacttt	gaccgctcct	gtgctcatga	ttgggtacat	cgactactgt	14340
ctgtgctaga	aaccacttta	ggagaaaagc	aagttttgcc	agcacgcaaa	ctcaggagca	14400
tggagggaatt	cacaaaaagg	tttccagatg	tgaaggaggt	gattgtggat	ggtagcgagc	14460
gtccagtcga	gcgtcctcaa	aaccgagaa	gcaaaaaaga	gtattactct	ggcaagaaaa	14520
agcggcctac	atgcaagcag	attacagtca	gcacaaggga	gaaacgagtg	attattcgga	14580
cggaaaccag	agcaggtaaa	gtgcatgaca	aacggctact	ccatgaatca	gagatagtgc	14640
aatacattcc	tgatgaagta	gcaatagagg	gagatttggg	ttttcatggg	ttggagaaag	14700
aatttgcata	tgtccattta	ccacacaaga	aaccgaaagg	tatcgaagca	aggaggcatg	14760
gcggcgggat	gggtcagttt	ttataagaga	gttttgacaa	tataaataaa	agacttttga	14820
caaccagact	tggcattact	tagtttcagt	ctttcatctc	aagtttacgt	tattctgagg	14880
cgaacatgaa	tcttataaca	acaaaaaac	aggtagatac	attagtatac	cacgctcatc	14940
tttttaccat	gcagggaaat	ggtgtgggat	atatgtcaga	tggggcactt	gcggttgagg	15000
gtagccgtat	tgtagcagtt	gattcgacgg	aggcggtgct	gagtcatttt	gagggcagaa	15060
agggttattga	gtccgcgaat	tgtgccgtct	tgccgtgggt	gattaatgct	cacgtagaca	15120
caagtttggt	gctgatgcgt	ggggcgggcg	aagatgtaac	taattggcta	atggacgcga	15180
ccatgcctta	ttttgctcac	atgacacccg	tgggcgagat	ggctgcaaca	cgcttaaggg	15240
tggtagaaga	gttgaagcga	ggcacaacaa	cattctgtga	caataaaatt	attagccccc	15300
tgtggggcga	atttttcgat	gaaattgggt	tacgggctag	tttagctcct	atgttcgatg	15360
cactcccact	ggagatgcca	ccgcttcaag	acggggagct	ttatcccttc	gatatacaag	15420
cgggacggcg	ggcgatggca	gaggctgtgg	attttgccctg	tggttggaat	ggggcagcag	15480
aggggcgtat	cactaccatg	ttaggaatgt	attcgccaga	tatgatgcgc	cttgagatgc	15540
tacgcgcagc	cagcagctac	aataatattt	aaggcttaatt	gctgcatttt	catgtagcgc	15600
agggagatgc	ggaaacagag	caaatcggtg	aacgatattg	taagcgctcg	atcgcatctc	15660
tagctgagat	tggtactctg	gacgaacagt	tgctggcagt	tcacctcacc	gatgccaccg	15720
atgaagaggt	gatacaagta	gcaaaaagtg	gcgctggcat	ggtactctgt	tcgggaatga	15780
ttggcactat	tgacgggtac	gtgccgcccg	ctcatgtgtt	tcggcaagca	ggcggaccgc	15840
ttgcgctagg	cagcagctac	aataatattt	tccatgagat	gaagctgacc	gccttattca	15900
acaaaataaa	atatcacgat	ccaaccatta	tgccggcttg	ggaagtctct	cgataggcta	15960
ccatcgaagg	agcgcgggcg	attgggtttag	atcacaagat	tggtctctct	gaagttggca	16020
aagaagccga	cctgatctta	atagacctca	gcacccctaa	cctctcacc	actctgctta	16080
accccatctg	taaccttgta	cctaatttcg	tgtacgctgc	ttcaggacat	gaagttaaaa	16140
gtgtcatggt	ggcgggaaaa	ctgttattgg	aagactacca	agtcttcaca	gtagatgagt	16200
ctgctatcat	tgtgaaagca	caattgcaag	cccaacagat	ttctcaatgc	gtagcatctg	16260
accctatcca	caaaaaaatg	gtgctgatgg	cggcgatggc	aaggggcca	ttgtaggaa	16320
ggtcttgagt	tatctagtaa	gctaagttgc	caactaaca	ttaaaaatac	gaagcaggtg	16380
ataaggcaga	attacagcag	gttgtctttc	ggatcgctcg	ttggatcttt	gtaccttccc	16440
tagtcatggc	gatcgccctc	atcgtcttcg	cccaaccctg	gatgagcctg	ttcggtgcag	16500
agtttgctgt	ggctcattgg	tagccgatac	catccctcca	actgacttgt	catgatagtc	16560
atggtgcgac	tttcccttcg	gtactgataa	actgggattg	aatccctttc	agagtcatca	16620
tgatagattt	gggaagtcta	aatgtggctg	agaagaaagt	gcttttccca	tgttgagaa	16680
agtcacatta	acatcagcat	caaaacgcct	aattctagat	tttacctatg	gtttcagcca	16740
aggtaaagga	actgagctca	aattacacgc	cgtcatgaga	taatatgatt	attaattttc	16800
tgtatagccc	agtttaattat	acttgattgt	aggctatttt	tagcctcttc	taatgaagaa	16860
tccagactaa	tcttatgtga	cgggaatatg	ttatgcaaga	aaaacgaatc	gcaatgtggt	16920
ctgtgccacg	aagtttggtg	acagtgtctg	tacaagcctg	gtcgagtcgg	ccagataaccg	16980

tagtctttga	tgaacttctc	tcctttccct	atctctttat	caaagggaaa	gatatgggct	17040
ttacttggac	agaccttgat	tctagccaaa	tgccccacgc	agattggcga	tccgtcatcg	17100
atctgttaaa	ggctccoctg	cctgaaggga	aatcaatcat	cgatctgtta	aaggctcccc	17160
tgctgaagg	gaaatcaatt	tgctatcaga	agcatcaagc	gtatcattta	atcgaagaga	17220
ccatggggat	tgagtggata	ttgcccttca	gcaactgctt	tctgattcgc	caacccaaag	17280
aaatgctctt	atcttttctg	aagattgtgc	cacattttac	ctttgaagaa	acaggctgga	17340
tcgaattaaa	acggctgttt	gactatgtac	atcaaacgag	cggagtaatc	ccgcctgtca	17400
tagatgcaca	cgacttgctg	aacgatccgc	ggagaatgct	ctccaagctt	tgtcagggtg	17460
taggggttga	gtttaccgag	acaatgctca	gttggccccc	catggagggtc	gagttgaacg	17520
aaaaactagc	cccttgggtac	agcaccgtag	caagttctac	gcattttcac	tcgtatcaga	17580
ataaaaaatga	gtcgttgccg	ctatatcttg	tcgatatattg	taaacgctgc	gatgaaatat	17640
atcaggaatt	atatcaattt	cgactttatt	agagagtatt	ggtaatgaaa	attttgaatt	17700
agtgaagaaa	tagaagttaga	gaatatagac	catctaggga	tagagactta	tgctggacgg	17760
attcaacaac	atcaggacaa	ttaccacagt	cagagtgtat	ttagctttgc	tgtttacgga	17820
caattatgga	tttatggcat	ggaactatag	gctgatttag	ctctaagctt	aattagtctt	17880
aaaectcata	aacgcctctt	tttcaagcgt	ggctttcagg	ctctatccct	tatgaaacaa	17940
gctgtttgac	cactttgtca	cccggtaagg	agaaaaacct	taaacccaag	cagaaaaaat	18000
tagcccgtaa	aaaaaaggga	agtaaataca	ggaaatatag	ggtaatatat	ttttcacaag	18060
tttatcaatt	gtaatctact	tgattcagta	aattaattaa	ggtgttgaa	agatgcaaac	18120
aagaattgta	aatagctgga	atgagtggga	tgaactaaag	gagatggttg	tcgggattgc	18180
agatgtgtct	tattttgaac	caactgagcc	cctgctttac	cggtataagaa	gcgataagaa	18240
cattgccaaa	atgttctctt	ttcccagggg	tccgaaaaag	caagaggtaa	cagagaaagc	18300
taatgaggag	ttgaatgggc	tggtagcgct	tctagaatca	cagggcgtaa	ctgtacgccg	18360
cccagagaaa	cataactttg	gcctgtctgt	gaagacacca	ttctttgagg	tagagaatca	18420
atattgtgcg	gtctgcccac	gtgatgttat	gatcaccttt	gggaacgaaa	ttctcgaagc	18480
aactatgtca	cgcgggtcac	gcttctttga	gtatttacct	tatcgcaaac	tagtctatga	18540
atattggcat	aaagatccag	atatgatctg	gaatgctgcg	cctaaaccga	ctatgcaaaa	18600
tgccatgtac	cgcgaagatt	tctgggagtg	tccgatggaa	gatcgatttg	agagtatgca	18660
tgattttgag	ttctgcgtca	cccaggatga	ggtgattttt	gacgcagcag	actgtagccg	18720
ctttggccgt	gatatttttg	tgcaggagtc	aatgacgact	aatcgtgcag	ggattcgctg	18780
gctcaaacgg	catttagagc	cgcgtcgctt	cgcgtgcat	gatattcact	tcccactaga	18840
tattttccca	tcccacattg	attgtacttt	tgtcccctta	gcacctgggg	ttgtgttagt	18900
gaatccagat	cgccccatca	aagagggtga	agagaaactc	ttcatggata	acggttggca	18960
attcatcgaa	gcacccctcc	ccacttccac	cgacgatgag	atgcctatgt	tctgccagtc	19020
cagtaagtgg	ttggcgatga	atgtgttaag	cattttcccc	aagaaggcca	tctgtgaaga	19080
gcaagagcat	cgccttcgat	agttgctaga	taaacacggc	tttgaggtct	atccaattcc	19140
ctttcgcaat	gtctttgagt	ttggcgggtc	gctccattgt	gccacctggg	atatccatcg	19200
cacgggaacc	tgtgaggatt	acttccctaa	actaaactat	acgccggtaa	ctgcatcaac	19260
caatggcggt	tctcgcttca	tcatttagta	ggttttatag	ttatgcaaaa	gagagaaagc	19320
ccacagatac	tatttgatgg	gaatggaaca	caatctgagt	ttccagatag	ttgcattcac	19380
cacttggtcg	aggatcaagc	cgcaaagcga	cgggatgcga	tcgctctcat	tgacggtgag	19440
caatccctta	cctacgggga	actaaatgta	cgcgctaacc	acctagccca	gcactctctg	19500
tccctaggct	gtcaaccgga	tgacctcctc	gccatctgca	tcgagcggtc	ggcagaactc	19560
tttatgggtt	tgttgggtat	cctaaaagcc	ggatgtgctt	atgtgccttt	ggatgtaggc	19620
tatctggcg	atcgcataga	gtatatgttg	cgggactcgg	atgcgcgtat	tttactaacc	19680
tcaacggatg	tcgctaagaa	acttgcccta	accatacctg	cattgcaaga	gtgccaacc	19740
gtctatttag	atcaagagat	atttgagtat	gattttcatt	ttttagcgat	agctaaacta	19800
ttacataacc	aatacttgag	attattacat	ttttattttt	ataccttgat	tcagcaatgc	19860
caggcaactt	cggtttccca	agggattcag	acacagggtc	tccccaataa	tctcgcttac	19920
tgcatttaca	cctctggctc	taccggaaat	cccaaaggga	tcttgatgga	acatcgctca	19980
ctggtgaata	tgctttgggtg	gcatcagcaa	acgcggcctt	cggttcaggg	tgtaggacg	20040
ctgcaatttt	gtgcagtcag	ctttgacttt	tcctgccatg	aaattttttc	taccctctgt	20100
cttggcggga	tattggtctt	ggtgccagag	gcagtgcgcc	aaaatccctt	tgcattggct	20160
gagttcatca	gtcaacagaa	aattgaaaaa	ttgtttcttc	ccgttatagc	attactacag	20220
ttggccgaag	ctgtaaattg	gaataaaaagc	acctccctcg	cgttttgcca	agttatcact	20280
accggggagc	agatgcagat	cacacctgct	gtgcgaacc	tctttcagaa	aaccggggcg	20340
atgttcata	atcactacgg	ggcaacagaa	tttcaagatg	ccaccactca	taccctcaag	20400

ggcaatccag	agggctggcc	aacactgggt	ccagtgggtc	gtccactgca	caatgttcaa	20460
gtgtatatcc	tggatgaggg	acagcaacct	gtacctcttg	gtggagaggg	tgaattctgt	20520
attggtggta	ttggactggc	tcgtggctat	cacaatttgc	ctgacctaac	gaatgaaaaa	20580
tttattccca	atccatttgg	ggctaattgag	aacgcataaa	aactctaccg	cacaggggac	20640
ttggcacgct	acctaccoga	cggcacgatt	gagcatttag	gacggataga	ccaccagggt	20700
aagatccgag	gtttccgcgt	ggaattgggg	gaaattgagt	ccgtgctggc	aagtcaccaa	20760
gctgtgcgtg	aatgtgccgt	tgtggcacgg	gagattgcag	gtcatacaca	gttggtaggg	20820
tatatcatag	caaaggatag	acttaatctc	agtttcgaca	aacttgaacc	tatcctgcgt	20880
caatattcgg	aagcgggtgt	gccagaatac	atgataccca	ctcggttcat	caatatcagt	20940
aatatgccgt	tgactcccag	tggtaaactt	gaccgcaggg	cattacctga	tcccaaaggc	21000
gatcgccctg	cattgtctac	cccacttgct	aagcctcgta	cccagacaga	gaaacgttta	21060
gcagagattt	ggggcagtta	tcttgctgta	gatattgtgg	gaacccacga	caatttcttt	21120
gatctaggcg	gtacgtcact	gctattgact	caagcgcaca	aattcctgtg	cgagaccttt	21180
aatattaatt	tgtccgcgtg	ctcactcttt	caatatccca	caattcagac	attggcacaa	21240
tatatgtatt	gccaaaggaga	cacaacctca	agcgatacag	catccaggca	caagaaagta	21300
cgtaaaaagc	agtcgggtga	cagcaacgat	attgccatca	tcagtgtggc	aggtcgcttt	21360
ccgggtgctg	aaacgattga	gcagttcttg	cataatctct	gtaatggtgt	tgaatccatc	21420
acctttttta	gtgatgatga	gctagagcag	actttgcctg	agttatttaa	taatcccgct	21480
tatgtcaaag	caggtgcggg	gctagaaggc	gttgaattat	ttgatgttac	cttttttggc	21540
tacagcccca	aagaagctgc	ggtgacagac	cctcagcaac	ggattttgct	agagtgtgcc	21600
tgggaagcat	ttgaacgggg	tggtctacaac	cccgaacct	atccagaacc	agttggtgtt	21660
tatgctgggt	caagcctgag	tacctatctg	cttaacaata	ttggtctctg	tttaggcata	21720
attaccgagc	aaccttttat	tgaaccggt	atggagcagt	ttcaggctaa	aattggcaat	21780
gaccggagct	atcttgcctac	acgcattctc	tacaagctga	atctcaaggg	tccaagcgct	21840
aatgtgcaga	ccgcctgctc	aacctcgcta	gttgcggttc	acatggcctg	tcagagtctc	21900
attagtggag	agtgctcaaat	ggcttttagcc	ggtggtat	ctgtggttgt	accacagaag	21960
gggggctatc	tctacgaaga	agggcatggt	cgttcccagg	atggctattg	tcgcgccttt	22020
gatgccgaag	cccaagggac	tatatgtggc	aatggcggcg	gcttgggttt	gcttaaaccg	22080
ttgcaggatg	cactggacga	taacgacaac	attatggcag	tcatacaagc	cacagccatc	22140
aacaacgacg	gtgcgctcaa	gatgggctac	acagaccoga	gcgtggatgg	gcaagctgat	22200
gtaattagcg	aggcgattgc	tatcgctgac	atagatgcaa	gcaccattgg	ctatgtagaa	22260
gctcatggca	cagccaccca	attgggtgat	ccgattgaag	tagcagggtt	agcaagggca	22320
tttcagcgta	gtacggacag	cgctcctggg	aaacaacaat	gcgctattgg	atcagttaaa	22380
actaatattg	gccacttaga	tgaggcggca	ggcattgccg	gactgataaa	ggctgctcta	22440
gctctacaat	atggacagat	tccaccgagc	ttgcactatg	ccaatcctaa	tccacggatt	22500
gattttgacg	caacccccatt	ttttgtcaac	acagaactac	gcgaatgggt	aaggaaatgg	22560
tatcctcggc	gggcgggggt	gagttctttt	ggtgtgggtg	gaactaacag	ccatattgtg	22620
ctggaggagt	cgctgttaaa	gcaaccacaa	ttgtctctct	ctttgccaga	acgcagtcac	22680
catctgctga	cgctttctgc	ccatacaca	gaggctttgc	atgagttggg	gcaacgctac	22740
atccaacata	acgagacaca	ccttgatatt	aacttagggc	acctctgttt	cacagccaat	22800
acgggacgca	agcattttga	gcctgcctta	gcggttgtag	ccgaatcaat	ccctggctta	22860
caggcacaca	tggaactgcg	acagactgcg	atttcagcac	agaaaaaaa	tgccccgcgg	22920
acgatcgcat	tcctgtttac	aggtcaaggc	tcacaataca	ttaacatggg	gcgcaccctc	22980
tacgatactg	aatcaacatt	ccgtgcagcc	cttgaccgat	gtgaaaccat	tctccaaaat	23040
ttagggatcg	agtccattct	ctccgttatt	tttggttcat	ctgagcatgg	actctcatta	23100
gatgacacag	cctataccca	gcccgcactc	tttgccatcg	aatacgcgct	ctatcaatta	23160
tggaagtcgt	ggggcatcca	gccctcagtg	gtgataggte	atagtgtagg	tgaatatgtg	23220
tccgcttggt	tggcgggagt	cttttagctta	gaggatgggt	tgaactgat	tgcagaacga	23280
ggacgactga	tacaggcact	tcctcgtgat	gggagcatgg	tttccgtgat	ggcaagcgag	23340
aagcgtattg	cagatatcat	tttaccttat	gggggacagg	tagggatcgc	cgcgattaat	23400
ggcccacaaa	gtgttgtaat	ttctgggcaa	cagcaagcga	ttgatgctat	ttgtgccatc	23460
ttggaactag	agggcatcaa	aagcaagaag	ctaaacgtct	cccatgcctt	ccactcgccg	23520
ctagtgggaag	caatgttaga	ctctttcttg	caggttgcac	aagaggtcac	ttactcgcaa	23580
cctcaaatca	agcttatctc	taatgtaacg	ggaacattgg	caagccatga	atcttgtccc	23640
gatgaaactc	cgatcaccac	cgcagagtat	tgggtacgtc	atgtgcgaca	gcccgtccgg	23700
tttgcggcgg	gaatggagag	ccttgagggt	caaggggtaa	acgtatttat	agaaatcggt	23760
cctaaacctg	ttcttttagg	catgggacgc	gactgcttgc	ctgaacaaga	gggactttgg	23820

ttgcctagtt	tgcgccccaa	acaggatgat	tggcaacagg	tgttaagtag	tttgcgtgat	23880
ctatacttag	cagggtgaac	cgtagattgg	agcagtttcg	atcaggggta	tgctcgtcgc	23940
cgtgtgccac	taccgactta	tcccttggcag	cgagagcggc	attgggtaga	gccaattatt	24000
cgtcaacggc	aatcagtatt	acaagccaca	aataccacca	agctaactcg	taacgccagc	24060
gtggcgagc	atcctctgct	tggccaacgg	ctgcatttgg	cgcggaactca	agagatttac	24120
tttcaaacct	tcatccactc	cgacttccca	atatgggttg	ctgatcataa	agtatttgga	24180
aatgtcatca	ttccgggtgt	cgctattttt	gagatggcac	tggcagcagg	gaaggcactt	24240
aaaccagaca	gtatatTTTT	gctcgaagat	gtatccatcg	cccaagcact	gattattccc	24300
gatgaagggc	aaactgtgca	aatagtatta	agcccacagg	aagagtcagc	ttattttttt	24360
gaaatcctct	ctttagaaaa	agaaaactct	tgggtgcttc	atgcctctgg	taagctagtc	24420
gccaagagc	aagtgcctaga	aaccgagcca	attgacttga	ttgcgttaca	ggcacattgt	24480
tccgaagaag	tgtcagtaga	tgtgctatat	caggaagaaa	tggcgcgccg	gctggatatg	24540
gggtccaatga	tgcgtggggg	gaagcagctt	tggcggtatc	cgctctcctt	tgccaaaagt	24600
catgatgcga	tcgcactcgc	caagggtcagc	ttgccagaaa	tcttgcttca	tgagtccaat	24660
gcctaccaat	tccatcctgt	aatcttggat	gcggggctgc	aaatgataac	ggtctcttat	24720
cctgaagcaa	accaaggcca	gacttatgta	cctgttggtg	tagagggctc	acaagttctat	24780
ggctgcctcca	gttcagaact	ttgggtgctgc	gcccataatc	ggcctccttt	ggatacagat	24840
caaaggcagg	gtattgattt	gctgccaaa	aaattgattg	cagacttgca	tctatttgat	24900
accagggttc	gtgtggttgc	catcatgttt	gggtgtgcaat	ctgtccttgt	gggacgggaa	24960
gcaatgttgc	gatcgcaaga	tacttggcga	aattggcttt	atcaagtctc	gtggaaacct	25020
caagcctgtt	ttggactttt	accgaattac	ctgccaaacc	cagataagat	tcggaaacgc	25080
ctggaaacaa	agttagcgac	attgatcatc	gaagctaatt	tggcgactta	tgcatcgcgc	25140
tatacccaac	tggaaagggt	aagtctagct	tacgttgtgg	cggttttccg	acaaatgggc	25200
tggctgtttc	aaccgggtga	gcgtttttcc	accgccaga	aggtatcagc	gttaggaatc	25260
gttgatcaac	atcggaact	attcgctcgt	ttgctcgaca	ttctagccga	agcagacata	25320
ctcgcgacgc	aaaacttgat	gacgatatgg	gaagtcattt	catacccgga	aacgattgat	25380
atacaggtag	ttcttgacga	cctcgaagcc	aaagaagcag	aagccgaagt	cacactgggt	25440
tcccgttgca	gtgcaaaatt	ggccgaagta	ttacaaggaa	aatgtgacct	catacagttg	25500
ctctttcccg	caggggacac	aacaacgtta	agcaaaactc	atcgtgaagc	cccagttttg	25560
ggtgttacta	atactctagt	ccaagaagcg	cttctttccg	ccctggagca	gttgccgccg	25620
gaacgtgggt	ggcgaatttt	agagattggt	gctggaacag	gtggaaccac	agcctacttg	25680
ttaccgcatc	tgcttgggga	tcagacaaaa	tatgtcttta	ccgatattag	tgcttttttt	25740
cttgccaaa	cggaagagcg	ttttaagat	taccggtttg	tacgttatca	ggtattagat	25800
atcgaaacag	caccacaggc	gcaaggattt	gaaccctaaa	tatacgattt	aatcgtagca	25860
gcggatgtct	tgcatgtctc	tagtgacctg	cgtaaaactc	ttgtacatat	ccggcaatta	25920
ttagcgcggg	gcgggatggt	gatcctgatg	gaagacagcg	aaccgcgacg	ctgggctgat	25980
ttaacctttg	gcttaacaga	aggctgggtg	aagtttacag	accatgactt	acgccccaac	26040
catccgctat	tgtctcctga	gcagtggcaa	atcttgttgt	cagaaatggg	atttagtcaa	26100
acaaccgcct	tatggccaaa	aatagatagc	ccccataaat	tgccacggga	ggcgggtgatt	26160
gtggcgcgta	atgaaccagc	catcagaaaa	ccccgaagat	ggctgatctt	ggctgacgag	26220
gagattgggt	gactactagc	caaacagcta	cgtgaagaag	gagaagattg	tatactcctc	26280
ttgccagggg	aaaagtacac	agagagagat	tcacaaacgt	ttacaatcaa	tcctggagat	26340
attgaagagt	ggcaacagtt	attgaaccga	gtaccgaaca	tacaagaaat	tgtacattgt	26400
tggagtatgg	tttccactga	cttagataga	gccactattt	tcagttgcag	cagtacgctg	26460
catttagttc	aagcattagc	aaactatcca	aaaaaccctc	gcttgtcact	tgtcacctca	26520
ggcgcacaa	ccgttaacga	acatcatggt	caaaatgtag	ttggagcagc	cctctggggc	26580
atgggaaaag	taattgcact	cgaacaccca	gagctacaag	tagcacaat	ggatttagac	26640
ccgaatggga	aggttaaggc	gcaagtagaa	gtgcttaggg	atgaacttct	cgccagaaaa	26700
gaccctgcat	cagcaatgtc	tgtgcctgat	ctgcaaacac	gacctcatga	aaagcaaata	26760
gcctttcgtg	agcaaacacg	ttatgtggca	agactttcgc	ccttagaccg	ccccaatcct	26820
ggagagaaa	gcacacaaga	ggctcttacc	ttccgtgatg	atggcagcta	tctgattgct	26880
ggtggttttag	gcggactggg	gttagtggtg	gctcgttttc	tggttacaaa	tggggctaaa	26940
taccttgtgc	tagtcggacg	acgtgggtcg	agggaggaac	agcaagctca	attaagcgaa	27000
ctagagcaac	tcggagcttc	cgtgaaagtt	ttacaagccg	atattgctga	tgcaagaaca	27060
ctagcccaag	cactttcagc	agtaacctac	ccaccattac	ggggtgttat	tcatgcggca	27120
ggtacattga	acgatgggat	tctacagcag	caaagttggc	aagcctttaa	agaagtgatg	27180
aatcccaagg	tagcagggtgc	gtggaacctc	catatactga	caaaaaatca	gccttttagc	27240

ttctttgtcc	tgttctcctc	cgccacctct	ttgttaggta	acgctggaca	agccaatcac	27300
gccgcccga	atgctttcct	tgatgggtta	gcctcctatc	gtcgtcactt	aggactaccg	27360
agcctctcga	ttaattgggg	gacatggagc	gaagtgggaa	ttgcggctcg	acttgaacta	27420
gataagttgt	ccagcaaaaca	gggagaggga	accattacgc	taggacaggg	cttaciaaatt	27480
cttgagcagt	tgctcaaaga	cgagaatggg	gtgtatcaag	tgggtgtcat	gcctatcaac	27540
tggacacaat	tcttagcaag	gcaattgact	ccgcagccgt	tcttcagcga	tgccatgaag	27600
agtattgaca	cctctgttag	taaactaacc	ttgcaggagc	gggactcttg	cccccaaggt	27660
tacgggcata	atattcgaga	gcaattagag	aacgctccgc	ccaaagaggg	tctgactctc	27720
ttgcaggctc	atgttcggga	gcagggtttcc	caagtttttg	ggatagacac	gaagacatta	27780
ttggcagaac	aagacgtggg	tttcttttacc	ctggggatgg	attcgtcgac	ctctgtcgag	27840
ttaagaaaca	ggttacaagc	cagtttgggc	tgctctcttt	cttccacttt	ggcttttgac	27900
tatccaacac	aacaggctct	tgtgaattat	cttgccaatg	aattgctggg	aacccttgag	27960
cagctacaag	agcctgaatc	tgatgaagaa	gatcagatat	cgtcaatgga	tgacatcgtg	28020
cagttgtctg	ccgcgaaact	agagatggaa	atttaagccc	atggatgaaa	aactaagaac	28080
atacgaacga	ttaatcaagc	aatcctatca	caagatagag	gctctggaag	ctgaagttaa	28140
caggttgaag	caaacccaat	gtgaacctat	cgccatcgtc	ggcatgggct	gtcgttttcc	28200
tgggtcgcaat	agtccagaag	cgttttggca	gttgttgtgt	gatgggggtg	atgctattcg	28260
tgagatacca	aaaaatcgat	gggttgttga	tgccctacata	gatgaaaatt	tggaccgcgc	28320
agacaagaca	tcaatgcgat	ttggcggggt	tgctcgagcaa	cttgagaagt	ttgatgccca	28380
attctttggc	atatcacgcg	gagaagcggg	ttctcttgac	cctcagcaac	gtttgttatt	28440
agaagttaagt	tgggaagcac	tggaataatgc	agcgttgata	ccaccttcgg	caacgggcgt	28500
attcgtcggt	attagtaacc	ttgattatcg	tgaacgcgtc	ttgaagcaag	gagcaattgg	28560
tacttatttt	gcttcgggta	atgcccatag	cacagccagt	ggtcgcttgt	cttactttct	28620
cggctctgaca	ggccccctgtc	tctcgataga	tacagcttgt	tcttcgtcgt	tggctcgtgt	28680
acatcagtca	ctgataagtc	tgcgctcagcg	agaattgtgac	ttagcgttgg	ttgggggagt	28740
ccatcggctg	atagccccag	aggaaaagtgt	ctcgttagca	aaagccata	tgttatctcc	28800
cgatgtctctg	tgcaaaagtct	ttgatgcgtc	ggcaaacggg	tatgtccgag	ccgaaggatg	28860
tggcatgata	gtcctcaaac	gattatcgga	cgcgcaagct	gatggggata	aaatcttggc	28920
gttgattcgc	gggtcagcca	taaatcaaga	cggctgcacg	agtggcttga	ccgttccaaa	28980
tgggtcccaa	caagccgacg	tgattcgcca	agcctcgcgc	aatagtggca	taagaccaga	29040
acaagttaac	tatgtagaag	ctcatggcac	agggacttcc	ctaggagacc	cgtattgaggt	29100
cggcgcgttg	ggaacgatct	ttaatcaacg	ctcccaacct	ttaattattg	gttcagttaa	29160
aacaaatatt	gggcattctag	aagcagcagc	agggattgct	ggactgatta	aagtcgtcct	29220
tgccatgcag	catggagaaa	ttccacctaa	tttacacttt	caccagccca	atcctcgcat	29280
taactgggat	aaattgccaa	tcaggatccc	cacagaacga	acagcttggc	ctactggcga	29340
tcgcatcgca	gggataagtt	ctttcggctt	tagtggcact	aattctcatg	tcgtgttaga	29400
ggaagcccca	aaaatagagc	cgtctacttt	agagattcat	tcaaagcagt	atgtttttac	29460
cttatcagca	gcgacacctc	aagcactaca	agaacttact	cagcgttatg	taacttatct	29520
cactgaacac	ttacaagaga	gtctggcgga	tatttgcctt	acagccaaca	cagggcgcaa	29580
acacttttaga	catcgctttg	cagtagtagc	agagtctaaa	acccagttgc	gccaacaatt	29640
ggaaacgttt	gccccaatcg	gagaggggca	ggggaagagg	acatctctct	caaaaatagc	29700
ttttctcttt	acagggtcaag	gctcacagta	tgtggggatg	gggcaagaac	tttatgagag	29760
ccaaccacc	ttccggcaaa	ccattgaccg	atgtgatgag	attcttcggt	cactgttggg	29820
caaatcaatc	ctctcaatac	tctatcccag	ccaacaaatg	ggattggaaa	cgccatccca	29880
aattgatgaa	accgcctata	ctcaaccac	tcttttttct	cttgaatatg	cactggcgca	29940
gttggtggcg	tcttggggta	ttgagcctga	tgtggtgatg	gggcatagtg	tgggagaata	30000
tgtggccgct	tgtgtggcgg	gtgtcttttc	tttagaggat	ggactcaaac	taattgctga	30060
aagaggccgt	ctgatgcaag	aattgcctcc	cgatggggcg	atggtttcag	ttatggccaa	30120
taaatcgcg	atagagcaag	caattcaatc	tgctagccga	gagggtttcta	ttgcggccat	30180
caatggacct	gagagtgtgg	ttatctctgg	taaaagggag	atattacaac	agattaccga	30240
acatctgggt	gccgaaggca	ttaagacacg	ccaactgaag	gtctctcatg	cctttcactc	30300
accattgatg	gagccaatat	taggtcagtt	ccgccgagtt	gccaatacca	tcacctatcg	30360
gccaccgcaa	attaaccttg	tctcaaatgt	cacaggcgga	caggtgtata	aagaaatcgc	30420
tactcccgat	tattgggtga	gacatctgca	agagactgtc	cgttttgcgg	atgggggttaa	30480
ggtgttacat	gaacagaatg	tcaatttcat	gctcgaaatt	ggtcccaaac	ccacactgct	30540
gggcattggt	gagttacaaa	gttctgagaa	tccattttct	atgccaatga	tgatgcccg	30600
tttgcgtcag	aatcgtagcg	actggcagca	gatgttggag	agcttgagtc	aactctatgt	30660

tcatggtgtt	gagattgact	ggatcggttt	taataaagac	tatgtgagac	ataaagttgt	30720
cctgccgaca	tacccatggc	agaaggagcg	ttactgggta	gaattggatc	aacagaagca	30780
cgccgctaaa	aatctacatc	ctctactgga	cagggtgcatg	aagctgcctc	gtcataacga	30840
aacaattttt	gagaaagaat	ttagtctaga	gacattgccc	tttcttgctg	actatcgcat	30900
ttatgggttca	gttgtgtcgc	cagggtgcaag	ttatctatca	atgatactaa	gtattgccga	30960
gtcgtatgca	aatggtcatt	tgaatggagg	gaatagtgcga	aagcaaacca	cttatttact	31020
aaaggatgtc	acattccag	tacctcttgt	gatctctgat	gaggcaaatt	acatgggtgca	31080
agttgcttgt	tctctctctt	gtgctgcgcc	acacaatcgt	ggcgacgaga	cgagtttga	31140
attgttcagt	tttgctgaga	atgtacctga	aagtagcagt	ataaatgctg	atttttcagac	31200
accatttatt	catgcaaaaag	ggcaatttaa	gcttgaagat	acagcacctc	ctaaagtggga	31260
gctagaagaa	ctacaagcgg	gttgtcccca	agaaattgat	ctcaaccttt	tctatcaaac	31320
attcacagac	aaaggttttg	tttttggatc	tcgttttctgc	tggttagaac	aaatctgggt	31380
ggcgcatgga	gaagcattgg	cgctctgctg	acaaccggaa	agtattgaat	cgtttaaagg	31440
atatgtgatt	catcccggtt	tggtggatgc	ctgtacacaa	gtcccatthg	caatttcgctc	31500
tgacgatgaa	aataggcaat	cagaaacgac	aatgcccttt	gcgctgaatg	aattacgttg	31560
ttatcagcct	gcaaacggac	aaatgtgggtg	ggttcatgca	acagaaaaag	atagatatatac	31620
atgggatgtt	tctctgtttg	atgagagcgg	gcaagttatt	gcggaattta	taggtttaga	31680
agttcgtgct	gctatgcccc	aaggcttact	aagggcagac	ttttggcata	actggctcta	31740
tacagtgaat	tgccgatcgc	aacctctaca	aatcccagag	gtgctggata	ttaataagac	31800
aggtgcagaa	acatggcttc	tttttgcaca	accagaggga	ataggagcgg	acttagccga	31860
atatttgcag	agccaaggaa	agcactgtgt	ttttgtagtg	cctgggagtg	agtatacagt	31920
gaccgagcaa	cacattggac	gcactggaca	tcttgatgtg	acgaaactga	caaaaattgt	31980
cacgatcaat	cctgcttctc	ctcatgacta	taaatattht	ttagaaactc	tgacggacat	32040
tagattacct	tgtagaacct	tactctattt	atggaatcgt	tatgatttaa	caaatacttc	32100
taatcatcgg	acagaattga	ctgtaccaga	tatagctcta	aacttatgta	ctagtcttac	32160
ttatttggta	caagccctta	gccacatggg	tttttccccg	aaattatggc	taattacaca	32220
aaatagtcaa	gcggttggta	gtgacttagc	gaatttagaa	atcgaacaat	ccccattatg	32280
ggcattgggt	cgaagcatcc	gcgcgaaca	ccctgaattt	gattgccgtt	gttttagattt	32340
tgacacgctc	tcaaatatcg	caccactctt	gttgaagag	atgcaagcta	tagactatga	32400
atctcaaatt	gcttaccgac	aaggaaacgg	ctatgttgca	cgactaatc	gtaatcaatc	32460
agaatgtcac	gcaccgatcc	aaacaggaat	ccgtcctgat	ggcagctatt	tgattacagg	32520
tggattaggc	ggtctaggat	tgcaggtagc	actcgccctt	gcggacgctg	gagcaagaca	32580
cttgatcctc	aatagtccgc	gtggtacgg	ctccaaagaa	gccagtttaa	ttattgaccg	32640
actacgcca	gaggatgta	gggttgattt	gattgcggca	gatgtctctg	atgcggcaga	32700
tagcgaacga	ctcttagtag	aaagtcagcg	caagacctct	cttcgagggg	ttgtccatgt	32760
tgcgggagtc	ttggatgatg	gcacccctgt	ccaacaaaat	caagagcggt	ttgaaaaagt	32820
gatggcggt	aaggtagcgc	gagcttggca	tctggaccaa	cagagccaaa	ccctcgattt	32880
agatttcttt	gttgcgcttc	catctgttgc	gtcgtcata	gaagaaccag	gacaagccaa	32940
ttacgcccga	gcgaatgcgt	ttttggattc	attaatgtat	tatcgtcaca	taaagggatc	33000
taatagcttg	agtatcaact	ggggggcttg	ggcagaagtc	ggcatggcag	ccaatttatc	33060
atgggaacaa	cggggaatcg	cggcaatttc	tccaaagcaa	gggaggcata	ttctcgtcca	33120
acttattcaa	aaacttaatc	agcatacaat	cccccaagtt	gctgtacaac	cgaccaattg	33180
ggctgaatat	ctatcccatg	atggcgtgaa	tatgccattc	tatgaatatt	ttacacacca	33240
cttgcgtaac	gaaaaagaag	ccaaattgcg	gcaaacagca	ggcagcacct	cagagggaagt	33300
cagtcctcgg	caacagcttc	aaacactctc	agagaaagac	cgggatgccc	ttttgatgga	33360
acatcttcaa	aaaactgcga	tcagagttct	cggtttggca	tctaatacaa	aaattgatcc	33420
ctatcaggga	ttgatgaata	tgggactaga	ctctttgatg	gcggttgaat	ttcggaatca	33480
cttgatacgt	agtttagaac	gccctctgcc	agccactctg	ctctttaatt	gcccacact	33540
tgattcattg	catgattacc	tagtcgcaaa	aatgtttgat	gatgcccctc	agaaggcaga	33600
gcaaatggca	caaccaacaa	cactgacagc	acacagcata	tcaatagaat	ccaaaataga	33660
tgataacgaa	agcgtggatg	acattgcaca	aatgctggca	caagcactca	atatcgcttc	33720
tgagttagca	tgggcagccc	ttaacctttc	aaggtgacta	atcaatagac	ctcttgca	33780
attgtttctg	tggtagaata	agtgttttta	ggttttatgt	atatttgggt	gttgttgcca	33840
tagctacgct	gcgccgaagg	atcacaaatt	caaagatagg	cgtgtgattc	taacttttag	33900
cttaacgggt	gacaaggcgg	ctaaagagct	tgtttcataa	gggatagagc	ctgaaagccc	33960
cgttgaaaaa	agagggcgtt	atgaggcttg	agattgatta	aattcagagc	taaatcagcc	34020
cataattcca	taccataaat	ccatagttgt	ccgtagagac	caaagctaaa	atcactttga	34080

cggtgggtact	tgtcctgatg	ttgtttgaatc	ccacattcag	catgagtaaa	tatactcaaa	34140
atatttttcc	cagcagggtta	agtgttctaa	tcctaagtct	gatatcttat	ttttgataag	34200
ggacttaccg	cgtaaatagtt	aaatttttgt	atagcctaata	tttacttggt	ttaaggctct	34260
tttttgctct	tttgggtgaat	tattcaggat	aatcaaagat	gagtcagccc	aattatggca	34320
ttttgatgaa	aaatgctgtg	aacgaaataa	atagcctacg	atcgcaacta	gctgcggtag	34380
aagcccaaaa	aaatgagttct	attgccattg	ttgggtatgag	ttgccgtttt	ccaggcggtg	34440
caactactcc	agagcggtttt	tgggtattac	tgcgcgaggg	tatatcagcc	attacagaaa	34500
tccctgctga	tcgctgggat	gttgataaat	attatgatgc	tgacccca	tcgtccggtg	34560
aaatgcatac	tcgttacggc	ggttttctga	atgaagttga	tacatttgag	ccatcattct	34620
ttaatatattg	tgcccgtgaa	gccgttagca	tggatccaca	gcaacgcttg	ctacttgaag	34680
tcagttggga	agctctggaa	tccggtaata	ttgttcctgc	aactcttttt	gatagttcca	34740
ctggtgtatt	tatcgggtatt	ggtggttagca	actacaaatc	tttaatgatc	gaaaacagga	34800
gtcgggacgg	gaaaaccgat	ttgtatgagt	taagtggcac	tgatgtgagt	gttgctgccc	34860
gcaggatatac	ctatgtcctg	ggttttgatgg	gtcccagttt	tgtgattgat	acagcttgtt	34920
catcttcttt	ggctctcagtt	catcaagcct	gtcagagtct	gcgtcagaga	gaatgtgac	34980
tagcactagc	tgggtggagtc	ggttttactca	ttgatccaga	tgagatgatt	ggtctttctc	35040
aaggggggat	gctggcacct	gatggtagtt	gtaaaacatt	tgatgccaat	gcaaatggct	35100
atgtgcgagg	cgaaggttgt	gggatgattg	ttctaaaacg	tctctcggat	gcaacagccg	35160
atggggataa	tattctttgcc	atcattcgtg	ggtctatggt	taatcatgat	ggtcatagca	35220
gtgggtttaac	tgctccaaga	ggccccgcac	aagtctctgt	cattaagcaa	gccttagata	35280
gagcagggtat	tgcaccggat	gccgtaagtt	atttagaagc	ccatgggtaca	ggcacacccc	35340
ttgggtgatcc	tatcgagatg	gattcattga	acgaagtgtt	tggctcggaga	acagaaccac	35400
tttgggtcgg	ctcagtttaag	acaaatattg	gtcattttaga	agccgcgtcc	ggtattgcag	35460
gggtgattaa	ggttgtcttg	atgctaaaaa	acaagcagat	tctctctcac	ttgcatttca	35520
agacaccaaa	tccatatatt	gattggaaaa	atctcccggg	cgaattcccg	accacccttc	35580
atgcttggga	tgacaagaca	ttgaaggaca	gaaagcgaat	tgcaggggtt	agttctttta	35640
gttttcagtgg	tactaacgcc	cacattgtat	tatctgaagc	cccatctagc	gaactaat	35700
gtaatcatgc	ggcagtgga	agaccatggc	acttggttaac	ccttagtgct	aagaatgagg	35760
aagcgttggc	taacttggtt	gggctttatc	agtcatttat	ttctactact	gatgcaagtc	35820
ttgccgatat	atgctacact	gctaatacgg	cacgaaccca	tttttctcat	cgcttgctc	35880
tatcggctac	ttcacacatc	caaatagagg	ctcttttagc	cgcttataag	gaagggtcgg	35940
tgagtttgag	catcaatcaa	ggttgtgtcc	tttccaacag	tcgtgcgcgg	aagggtcgctt	36000
ttctctttac	aggtcaaggt	tcgcaatatg	tgcaaatggc	tggagaactt	tatgagaccc	36060
agcctaacttt	ccgtaattgc	ttagatcgct	gtgccgaaat	cttgcaatcc	atcttttcat	36120
cgagaaacag	cccttgggga	aaccactgc	tttcgggtatt	atatccaaac	catgagtcaa	36180
aggaaattga	ccagacggct	tatacccaac	ctgccctttt	tgtgttagaa	tatgccctag	36240
cacagatgtg	gcggctcgtg	ggaatcgagc	cagatatcgt	aatgggtcat	agcatagggtg	36300
aatatgtggc	agcttgtgtg	gcggggatct	tttctctgga	ggatgggtctc	aaacttgctg	36360
ccgaaaagagg	ccgtttgatg	caggcgctac	cacaaaatgg	cgagatgggt	gctatatcgg	36420
cctcccttga	ggaagttaag	ccggctattc	aatctgacca	gcgagttgtg	atagcggcgg	36480
taaatggacc	acgaagtgtc	gtcatttcgg	gcgatcgcca	agctgtgcaa	gtcttcacca	36540
acaccctaga	agatcaagga	atccggtgca	agagactgtc	tgtttcacac	gctttccact	36600
ctccattgat	gaaaccaatg	gagcaggagt	tcgcacaggt	ggccagggaa	atcaactata	36660
gtcctccaaa	aatagctctt	gtcagtaatc	taaccggcga	cttgatttca	cctgagtcct	36720
ccctggagga	aggagtgtac	gcttccccctg	gttactgggt	aaatcattta	tgcaatcctg	36780
tcttgttctg	tgatgggtatt	gcaactatgc	aagcgagga	tgtccaagtc	ttccttgaag	36840
ttggaccaaa	accgacctta	tcaggactag	tgcaacaata	ttttgacgag	gttgcccata	36900
gcgatcgccc	tgtoaccatt	cccaccttgc	gccccaaagca	acccaactgg	cagacactat	36960
tggagagttt	gggacaactg	tatgcgcttg	gtgtccaggt	aaattgggag	ggcttttgata	37020
gagattacac	cagacgcaaa	gtaagcctac	ccacctatgc	ttggaagcgt	caacgttatt	37080
ggctagagaa	acagtcgcgt	ccacgttttag	aaacaacaca	agttcgtccc	gcaactgcca	37140
ttgtagagca	tcttgaacaa	ggcaatgtgc	cgaaaatcgt	ggacttggtta	gcggcgacgg	37200
atgtactttc	agggcaagca	cggaatttgc	tacccagcat	cattgaacta	ttggttgcaa	37260
aacatcgtga	ggaagcgaca	cagaagccca	tctgcgattg	gctttatgaa	gtgggtttggc	37320
aaccccagtt	gctgacccta	tctaccttac	ctgctgtgga	aacagagggt	agacaatggc	37380
tcacttctgc	cgatgctagt	ggacacgggtg	aagcacttgc	ggctcaatta	cgtcagcaag	37440
gggatataat	tacgcttctg	tatgctgggtc	taaaatatca	ctcggctaat	aataaacaaa	37500

ataccggggg	ggacatccca	tatttttcaga	ttgatccgat	ccaaagggag	gattatgaaa	37560
ggttgtttgc	tgctttgcct	ccactgtatg	gtattgttca	tctttggagt	ttagatatac	37620
ttagcttggga	caaagtatct	aacctaattg	aaaatgtaca	attaggtagt	ggcacgctat	37680
taaatttaac	acagacagtc	ttgcaacttg	aaacgcccac	ccctagcttg	tggctcgtga	37740
caaagaacgc	gcaagctgtg	cgtaaaaaac	atagcctagt	cggagtgcct	cagtcaccct	37800
tatgggggat	gggtaagggt	atagccttag	aaacaccctga	actcaactgt	gtatcaatcg	37860
accttgatgg	tgaagggcct	ccagatgaac	aagccaagtt	tctggcggct	gaactccgcg	37920
ccgcctccga	gttcagacat	accaccattc	cccacgaaag	tcaagttgct	tggcgtaata	37980
ggactcgcta	tgtgtcacgg	ttcaaagggt	atcagaagca	tcccgcgacc	tcatcaaaaa	38040
tgccatttcg	accagatgcc	acttatttga	tcacgggcgg	ctttggtggt	ttgggcttgc	38100
ttgtggctcg	ttggatggtt	gaacaggggg	ctacccatct	atcttctgat	ggacgcagcc	38160
aacccaaccc	agccgcccac	aaacaactgc	aagagatagc	cgcgctgggt	gcaacagtga	38220
cgggtggtgca	agccgatggt	ggcatccgct	ccaagtagc	caatgtgttg	gcacagattg	38280
ataaggcata	tcctttggct	ggtattattc	atactgccgg	tgtattagac	gacggaatct	38340
tattgcagca	aaattgggcy	cgttttagca	aggtgttcgc	ccccaaacta	gagggagctt	38400
ggcatctaca	tacactgact	gaagagatgc	cgcttgattt	ctttatttgt	ttttcctcaa	38460
cagcaggatt	gctgggcagt	ggtggacaag	ctaactatgc	tgctgccaat	gccttttttag	38520
atgcctttgc	ccatcatcgg	cgaatacaag	gcttgccagc	tctctcgatt	aactgggacg	38580
cttgggtcga	agtgggaatg	acggtacgtc	tccaacaagc	ttcttcacaa	agcaccacag	38640
ttgggcaaga	tattagcatt	ttggaaaattt	caccagaaca	gggattgcaa	atctttgcct	38700
atcttctgca	acaaccatcc	gcccacaatag	cgccatttct	taccgatggg	cttcgcaaga	38760
tgtacgacac	aagctcggcc	ttttttgctt	tacttgatct	tgacagggtct	tcctccacta	38820
cccaggagca	atctacactt	tctcatgaag	ttggccttac	cttactcgaa	caattgcagc	38880
aagctcggcc	aaaagagcga	gagaaaatgt	tactgcgcca	tctacagacc	caagttgtcg	38940
cggctcttgcg	tagtcccga	ctgcccgcag	ttcatcaacc	cttcactgac	ttggggatgg	39000
attcgttgat	gtcacttgaa	ttgatgcggc	gtttggaaga	aagctcgggg	attcagatgc	39060
ctgcaacgct	tgcatctcat	tatcctatgg	tagaccgttt	ggctaagttt	atactgactc	39120
aaatatgtat	aaattctgag	ccagatacct	cagcagttct	cacaccagat	ggaaatgggg	39180
aggaaaaaga	cagtaataag	gacagaagta	ccagcacttc	cgttgactca	aattattact	39240
ccatggcaga	agattttatc	gcactcgaat	ccttactaaa	taaaaataaa	agagatcaat	39300
aatagagctg	ttgggaaata	aaagcatatt	tccggatgac	agaacttccc	ccatcccgat	39360
tgaatttatg	ctgcatctaa	atagaagttc	catagccctg	cactgaccaa	catcaattga	39420
tcatacaaat	cggtcacacg	attcctatat	gtgggataaa	atttgcagta	cagcaggata	39480
taaaatagtt	tttctcttat	acttctgagt	gtaggcttgc	gtcccgcctc	gggcgcacgt	39540
ttgcggtttg	ctaaggagtt	gaacacggtg	cgttcatagg	tatcagcaaa	ctgagataac	39600
agctcgttga	atgcttggcg	gttaagtcca	gtcattgctc	gtagcagtcg	ctcttgattc	39660
aggatgcggg	ctaagttcaa	cattaatgtc	accctacttg	tctgcttgat	tattatccct	39720
tattttccaa	caactctatt	atagcttate	ttatttttga	gtttaactac	atgaaaatcg	39780
ctgtaaaagc	tctactgag	tgaagtga	cttctttccc	acgtattcga	gtagctgttg	39840
taagctggcc	tcgatggaaa	gttccgaagt	ttccaccagt	aaatctgggt	ttctcgggtg	39900
ttcgtaggga	gcgctaattc	ccgtaaaaga	ctcaatttct	ccacggcggt	cttttgcata	39960
gagacccttg	gggtcacggt	gttcacaaat	ttccatcgga	gttgcaatat	atacttcatg	40020
aaacagatct	ccggacagaa	tacggatttg	ctcccgtctc	ttcctgtaag	gtgaaatgaa	40080
agcagtaatc	actaaacaac	ccgaatccgc	aaaaagtttg	gccacctcgc	caatacgacg	40140
aataattttcc	gcacgatcag	cagcagaaaa	tcccaagtca	gcacataatc	catgacggat	40200
attgtcacca	tcaaggacaa	aagtatacca	acctttcttg	aacaaaatcc	gctctaattc	40260
tagagccaat	gttgttttac	ctgatccctga	taatccagtg	aaccatagaa	ttccatttctg	40320
gtgaccattc	tttaaaacaac	gatcaaatgg	ggacacaaga	tgttttgtat	gttgaatatt	40380
gcttgatttc	atatctatga	taaatatgat	aaaagtgtat	ggccaaacag	aactgctcac	40440
ccaataatat	agttaaagggt	tatttttttca	aaaactcctt	ctaaattata	gctcacaatt	40500
atgcctaaat	actttaatac	tgctggaccc	tgtaaatccg	aaatccacta	tatgctctct	40560
cccacagctc	gactaccgga	tttgaagca	ctaattgacg	gagaaaaacta	ctttataatt	40620
cacgcgccgc	gactaagtcg	caaaaactaca	gctatgatag	ccttagcacg	agaattgact	40680
gatagtggaa	aataataccgc	agttatttctt	tccgttgaag	tgggatcagt	attctcccat	40740
aatccccagc	aagcggagca	ggttattttta	gaagaatgga	aacaggcaat	caaattttat	40800
ttacccaaag	aactacaacc	atcctatttg	ccagagcgtg	aaacagactc	aggaataggc	40860
aaaacttta	gtgagtggtc	cgcacaatct	ccaagacctc	ttgtaatctt	tttacatgaa	40920

atcgattccc	taacagatga	agctttaatc	ctaattttta	gacaattacg	ctcagggtttt	40980
ccccgtcgtc	ctcggggatt	tccccattcg	gtgggggttaa	ttggatgctg	ggatgtgcgg	41040
gactataaag	ttaaatctgg	tggaagtga	cgactgaata	cgtcaagtcc	tttcaatc	41100
aaagcggaat	ccttgacttt	aagtaatttc	actctgtcag	aggtggaaga	actttactta	41160
caacatacgc	aagctacagg	acaaaatttt	accccggaag	caattaaaca	agcattttat	41220
ttaaccgatg	ggcaaccatg	gttagtaaac	gccctagctc	gtcaagccac	tcagggtgta	41280
gtgaaagata	ttactcaacc	cattaccgct	gaagtaatta	accaagccaa	agaagttctg	41340
attcagcgcc	aggataccca	tttggatagt	ttggcagagc	gcttacggga	agatcgggtc	41400
aaagccatta	ttcaacctat	gttagctgga	tccgacttac	cagatacccc	agaggatgat	41460
cgccgtttct	tgctagattt	aggcttggtg	aagcgcgatc	ccttgggagg	actaaccatt	41520
gccaatccca	tttaccagga	ggtgattcct	cgtgttttgt	cccagggtag	tcaggatagt	41580
ctaccccgaga	ttcaacctac	ttggttaaat	actgataata	ctttaaatcc	tgacaaactc	41640
ttaaatgctt	tcctagagtt	ttggcgacaa	catggggaac	cattactcaa	aagtgcgcct	41700
tatcatgaaa	ttgctcccca	tttagttttg	atggcgcttt	tacatcgggt	agtgaatggt	41760
ggtggcactt	tagaacggga	atatgccgtt	ggttctggaa	gaatggatat	ttgtttacgc	41820
tatggcaagg	tagtgatggg	catagagtta	aagggttggg	ggggaaaatc	ggatccgtta	41880
acgaagggtt	tgacccaatt	ggataaatat	ctgggtgggt	taggattaga	tagagggttg	41940
ttagttaatt	ttgatcacgc	tccgggatta	ccacccatgg	gtgagaggat	tagtatggaa	42000
caggccatta	gtccagaggg	aagaaccatt	acagtgattc	gtagctagag	cgtagatat	42060
cagatgattg	aacctcaatt	attgtgcaac	gccacatttt	ctttccaaag	atgtatgtta	42120
aactctagta	aactctaatt	aggtcgagaa	agagat			42156

<210> 81

<211> 5631

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 81

atgcggtcta	agttcaacat	taatgtcacc	ctacttgtct	gcttgattat	tatcccttat	60
tttccaacaa	ctctaagtga	agtacctata	acagcaaacg	aagatgcagc	tacattactt	120
cagctgtgtg	gactgtccct	aaaggaagca	caccaacaac	ttgaggcaat	gcaacgccga	180
gcgcacgaac	cgatcgcaat	tgtggggctg	gggctgcggt	ttccgggagc	tgattcacca	240
cagacattct	ggaaactact	tcagaatggt	gttgatattg	tcaccgaaat	ccctagcgat	300
cgctgggcag	ttgatgaata	ctatgatccc	caacctgggt	gtccaggcaa	aatgtatatt	360
cgtgaagccg	cttttgttga	tgcatgtgat	aaattcgatg	cctcgttttt	tgatatttcg	420
ccacgtgaag	ggccaatat	agatccccag	catagaatgt	tgctggagggt	agcttgggag	480
gcactcgaaa	gggctggcat	tgctcccagc	caattgatgg	atagccaaac	gggggtattt	540
gtcgggatga	gcgaaaatga	ctattatgct	cacctagaaa	atcacaggga	tcatacataat	600
gtctatgcgg	caacgggcaa	tagcaattac	tatgctccgg	ggcggtttat	ctatctattg	660
gggcttcaag	gacctaacat	ggtcgttgat	agtgcctgtt	cctcctcctt	agtggctgta	720
catcttgctt	gtaatagttt	gcggatggga	gaatgtgata	tgccactggc	tggtggcggt	780
cagcttatgt	taatcccaga	ccctatgatt	gggactgccc	agttaaatgc	ctttgcgacc	840
gatggtcgta	gtaaaacatt	tgacgctgcc	gccgatggct	atggacgcgg	cgaaggttgt	900
ggcatgattg	tacttaaaaag	aataagtgac	gcgatcgtgg	cagacgatcc	aatttttagcc	960
gtaatccggg	gtagtgcagt	caatcatggc	gggcgttagca	gtggtttaac	tgcccctaata	1020
aagctgtctc	aagaagcctt	actgcgtcag	gcactacaaa	acgccaaagg	tcagccggaa	1080
gcagtcagtt	atatcgaagc	ccatggcaca	gggacacaa	tgggcgaccc	gattgagggtg	1140
ggagcattaa	cgaccgtctt	tgatctttct	cgttcagaac	ccttctggat	tggtctctgtc	1200
aaaactaata	tcggacacct	agaaccagcc	gctgggtatt	cggggttaata	aaaagtcatt	1260
ttatcattac	aagaaaaaca	gattcctccc	agtcctccatt	ttcaaaaacc	taatcccttc	1320
attgattggg	aatcttccgc	agttcaagtg	ccgacacagt	gtgtaccctg	gactgggaaa	1380
gagcgcgtcg	ctggagtttag	ctcgtttggt	atgagcggtg	caaactgtca	tctagtgtgc	1440
gcagaagcac	ctgtccgcca	aaacgaaaaa	tctgaaaatg	caccggagcg	tccttgtcac	1500
attctgaccc	tttcagccaa	aaccgaagcg	gcactcaacg	cattggtagc	ccgttacatg	1560
gcattttctca	gggaagcgcc	cgccatatcc	ctagctgata	tttgttatag	tgccaatgtc	1620
gggcgtaatc	tttttgccca	tcgcttaagt	tttatctccg	agaacatcgc	gcagttatca	1680
gaacaattag	aacactgccc	acagcaggct	acaatgccaa	cgcaacataa	tgtgatacta	1740
gataatcaac	tcagccctca	aatcgctttt	ctgtttactg	gacaagggtc	gcagtacatc	1800

10

aacatggggc	gtgagcttta	cgaaaactcag	cccaccttcc	gtcggattat	ggacgaatgt	1860
gacgacattc	tgcatccatt	gttgggtgaa	tcaattctga	acatactcta	cacttcccct	1920
agcaaaactta	atcaaaacgt	ttatacccaa	cctgcccttt	ttgcttttga	atatgcccta	1980
gcaaaactat	ggatatcatg	gggtattgag	cctgatgtcg	tactgggtca	cagcgtgggt	2040
gaatatgtag	ccgcttgtct	ggcgggtgtc	tttagtttag	aagatgggtt	aaaactcatt	2100
gcatctcgtg	gatgtttgat	gcaagcctta	cgcgcgggga	aaatgcttag	tatcagaagc	2160
aatgagatcg	gagtgaaagc	gctcatcgcg	ccttatagtg	cagaagtatc	aattgcagca	2220
atcaatggac	agcaaaagcg	ggtgatctcc	ggcaaaagctg	aaattataga	taatttagca	2280
gcagagttag	catcggaagg	catcaaaaaca	cacctaat	cagtctccca	cgctttccac	2340
tcgccaatga	tgacccccat	gctgaaagca	ttccgagacg	ttgccagcac	catcagctat	2400
aggtcaccca	gtttatcact	gatttctaac	ggtagacggc	aattggcaac	aaaggaggtt	2460
gctacacctg	attattgggt	gcgtcatgtc	cattctaccg	tccgttttgc	cgatgggtatt	2520
gccacattgg	cagaacagaa	tactgacatc	ctcctagaag	taggacccaa	accaatatgt	2580
ttgggtatgg	caaagcagat	ttatagtga	aacggttcag	ctagtcatcc	gctcatgtca	2640
cccagtttgc	gtgaagatgg	caacgattgg	cagcagatgc	tttctacttg	tggacaactt	2700
gtagttaatg	gagtcgaagt	tgactggcg	ggttttgaca	aggattattc	acgacacaaa	2760
atattgttgc	ccacctatcc	gtttcagaga	gaacgatatt	ggattgaaag	ctccgtcaaa	2820
aagccccaaa	aacaggagct	gcgcccgaatg	ttggataaga	tgatccggct	accatcagag	2880
aacaaagtgg	tggttgaaac	cgagtttggc	gtgcgacaga	tgctcatat	ctccgatcat	2940
cagatatacg	gtgaagtcat	tgtaccgggg	gcagtattag	cttcttaaat	cttcaatgca	3000
gcgcaggttt	tatacccaaga	ctatcagcat	gaattaaactg	atattgcttt	ttatcagcca	3060
attatctttc	atgacgacga	tacgggtgatc	gtgcaggcga	ttttcagccc	tgataagtca	3120
caggagaatc	aaagccatca	aacattttcca	cccatgagct	tccagattat	tagcttcatg	3180
ccggatggtc	ccttagagaa	caaaccgaaa	gtccatgtca	cagggtgtct	gagaatgttg	3240
cgcgatggcc	aaccgccaac	actctccccg	accgaaatac	gtcagcgctg	tccacatacc	3300
gtaaatggtc	atgactggta	caatagctta	ctcaaacaaa	aatttgaaat	gggtccttcc	3360
tttaggtggg	tacagcaact	ttggcatggg	gaaaatgaag	cattgaccgc	tcttcacata	3420
ccagatgtgg	tcggctctgt	atcaggacat	caacttcacg	gcatattgct	cgatgggtca	3480
ctttcaacca	ccgtctgcat	ggagtacgag	tacggagact	ccgcgaccag	agttcctttg	3540
tcatttgctt	ctctgcaact	gtacaaaccc	gtcacgggaa	cagagtgggtg	gtgctacgcg	3600
aggaatgattg	gttacttcaa	atatgacttc	cagattatga	atgaaatcgg	ggaaaccttg	3660
gtgaaagcaa	ttggctttgt	acttcgtgaa	gcctctcccg	aaaaattcct	cagaacaaca	3720
tacgtacaca	actggcttgt	agacattgaa	tggaagctc	aatcaacttc	cctagtcctt	3780
tctgatggca	ctatctcttg	cagttgtttg	gttttatcag	atcagcatgg	aacaggggct	3840
gcattggcac	aaaggctaga	caatgctgga	gtgccagtga	ccatgatcta	tgctgatctg	3900
atactggaca	attaccgaatt	aatattccgt	actttgccag	atttacaaca	agtcgtctat	3960
ttatgggggt	tggatcaaaa	agaggattgt	caccccatga	agcaagcaga	ggataactgt	4020
acatcggtgc	tatatcttgt	gcaagcatta	ctcaatacct	actcaacccc	gccatccctg	4080
cttattgtca	cctgtgatgc	acaagcggtg	gttgaaacaag	atcgagtaaa	tggcttcgcc	4140
caatcgctct	tggtgggact	tgccaaagtt	atcatgctag	aacacccaga	attgtcctgt	4200
gtttacatgg	atgtgggaagc	cggatattta	cagcaagatg	tggcgaacac	gatatttaca	4260
cagctaaaaa	gaggccatct	atcaaaggac	ggagaagaga	gtcagttggc	ttggcgcaat	4320
ggacaagcat	acgtagcacg	tcttagtcaa	tataaaccca	aatccgaaca	actgggtgag	4380
atccgcagcg	atcgcagcta	tttgatcact	ggtggacggg	gcggtgtcgg	cttacaatc	4440
gcacggtggg	tagtgaaaaa	gggggctaaa	catctcgttt	tgttggggcg	cagtcagacc	4500
agttccgaag	tacgtctggt	gttggaatgag	ctagaatcag	ccggggcgca	aatcattgtg	4560
gctcaagctg	atattagcga	tgagaaggta	ttagcgcaga	ttctgaccaa	tctaaccgta	4620
cctctgtgtg	gtgtaatcca	cgccgcagga	gtgcttgatg	atgcgagtct	actccaacaa	4680
actccagcca	agctcaaaaa	agttctattg	ccaaaagcag	agggggcttg	gattctgcat	4740
aatttgaccc	tggagcagcg	actagacttc	tttgttctct	tttcttctgc	cagttctcta	4800
ttagtgcgcg	cagggcaggc	caactattca	gcagccaatg	ctttcctaga	tggtttagct	4860
gcctatcgcg	gagggcgagg	actccctgt	ttgtctatct	gctggggggc	atgggatcaa	4920
gtcggtatgg	ctgcacgaca	agggctactg	gacaagttac	cgcaaagagg	tgaagaggcc	4980
atcccgttac	agaaaggctt	agacctcttc	ggcgaattac	tgaacgagcc	agccgctcaa	5040
attggtgtga	tcccaattca	atggactcgc	ttcttggatc	atcaaaaagg	taatttgctt	5100
ttttatgaga	agttttctaa	gtctagccgg	aaagcgcaga	gttacgattc	gatggcagtc	5160
agtcacacag	aagatattca	gaggaaactg	aagcaagctg	ctgtgcaaga	tcgacaaaa	5220
ttattagaag	tgcatcttcg	ctctcaagtc	gctcaactgt	taggaataaa	cgtggcagag	5280
ctaccaaagt	aagaaggaat	tggttttgtt	acattaggct	ttgactcgct	cacctctatt	5340
gaactgcga	acagtttaca	acgcacatta	gattgttcat	tacctgtcac	ctttgctttt	5400
gactacccaa	ctatagaaat	agcggtaag	tacctaacac	aagttgtaat	tgcaccgatg	5460
gaaagcacag	catcgcagca	aacagactct	ttatcagcaa	tgttcacaga	tacttcgtcc	5520
atcggggaga	ttcttgacaa	cgaacagat	gtgttagaca	gcgaaatgca	aagtgatgaa	5580
gatgaatctt	tgtctacact	tatacaaaaa	ttatcaacac	atttgatta	g	5631

- 5 <210> 82
 <211> 1876
 <212> PRT
 <213> *Cylindrospermopsis raciborskii* AWT205

<400> 82

Met	Arg	Ser	Lys	Phe	Asn	Ile	Asn	Val	Thr	Leu	Leu	Val	Cys	Leu	Ile
1				5					10					15	
Ile	Ile	Pro	Tyr	Phe	Pro	Thr	Thr	Leu	Met	Lys	Val	Pro	Ile	Thr	Ala
			20					25					30		
Asn	Glu	Asp	Ala	Ala	Thr	Leu	Leu	Gln	Arg	Val	Gly	Leu	Ser	Leu	Lys
		35						40				45			
Glu	Ala	His	Gln	Gln	Leu	Glu	Ala	Met	Gln	Arg	Arg	Ala	His	Glu	Pro
		50						55				60			
Ile	Ala	Ile	Val	Gly	Leu	Gly	Leu	Arg	Phe	Pro	Gly	Ala	Asp	Ser	Pro
65					70					75					80
Gln	Thr	Phe	Trp	Lys	Leu	Leu	Gln	Asn	Gly	Val	Asp	Met	Val	Thr	Glu
				85					90					95	
Ile	Pro	Ser	Asp	Arg	Trp	Ala	Val	Asp	Glu	Tyr	Tyr	Asp	Pro	Gln	Pro
			100					105					110		
Gly	Cys	Pro	Gly	Lys	Met	Tyr	Ile	Arg	Glu	Ala	Ala	Phe	Val	Asp	Ala
		115						120				125			
Val	Asp	Lys	Phe	Asp	Ala	Ser	Phe	Phe	Asp	Ile	Ser	Pro	Arg	Glu	Ala
		130					135					140			
Ala	Asn	Ile	Asp	Pro	Gln	His	Arg	Met	Leu	Leu	Glu	Val	Ala	Trp	Glu
145					150					155					160
Ala	Leu	Glu	Arg	Ala	Gly	Ile	Ala	Pro	Ser	Gln	Leu	Met	Asp	Ser	Gln
				165					170					175	
Thr	Gly	Val	Phe	Val	Gly	Met	Ser	Glu	Asn	Asp	Tyr	Tyr	Ala	His	Leu
			180					185					190		
Glu	Asn	Thr	Gly	Asp	His	His	Asn	Val	Tyr	Ala	Ala	Thr	Gly	Asn	Ser
		195						200					205		
Asn	Tyr	Tyr	Ala	Pro	Gly	Arg	Leu	Ser	Tyr	Leu	Leu	Gly	Leu	Gln	Gly
		210				215						220			
Pro	Asn	Met	Val	Val	Asp	Ser	Ala	Cys	Ser	Ser	Ser	Leu	Val	Ala	Val
225					230					235					240
His	Leu	Ala	Cys	Asn	Ser	Leu	Arg	Met	Gly	Glu	Cys	Asp	Leu	Ala	Leu
				245					250					255	
Ala	Gly	Gly	Val	Gln	Leu	Met	Leu	Ile	Pro	Asp	Pro	Met	Ile	Gly	Thr
			260					265					270		
Ala	Gln	Leu	Asn	Ala	Phe	Ala	Thr	Asp	Gly	Arg	Ser	Lys	Thr	Phe	Asp
		275						280					285		
Ala	Ala	Ala	Asp	Gly	Tyr	Gly	Arg	Gly	Glu	Gly	Cys	Gly	Met	Ile	Val
		290				295					300				
Leu	Lys	Arg	Ile	Ser	Asp	Ala	Ile	Val	Ala	Asp	Asp	Pro	Ile	Leu	Ala
305					310					315					320
Val	Ile	Arg	Gly	Ser	Ala	Val	Asn	His	Gly	Gly	Arg	Ser	Ser	Gly	Leu
				325					330					335	
Thr	Ala	Pro	Asn	Lys	Leu	Ser	Gln	Glu	Ala	Leu	Leu	Arg	Gln	Ala	Leu
			340					345					350		

Gln	Asn	Ala	Lys	Val	Gln	Pro	Glu	Ala	Val	Ser	Tyr	Ile	Glu	Ala	His
	355						360					365			
Gly	Thr	Gly	Thr	Gln	Leu	Gly	Asp	Pro	Ile	Glu	Val	Gly	Ala	Leu	Thr
	370					375					380				
Thr	Val	Phe	Gly	Ser	Ser	Arg	Ser	Glu	Pro	Leu	Trp	Ile	Gly	Ser	Val
385					390					395					400
Lys	Thr	Asn	Ile	Gly	His	Leu	Glu	Pro	Ala	Ala	Gly	Ile	Ala	Gly	Leu
				405					410					415	
Ile	Lys	Val	Ile	Leu	Ser	Leu	Gln	Glu	Lys	Gln	Ile	Pro	Pro	Ser	Leu
		420						425						430	
His	Phe	Gln	Asn	Pro	Asn	Pro	Phe	Ile	Asp	Trp	Glu	Ser	Ser	Pro	Val
	435						440						445		
Gln	Val	Pro	Thr	Gln	Cys	Val	Pro	Trp	Thr	Gly	Lys	Glu	Arg	Val	Ala
	450					455					460				
Gly	Val	Ser	Ser	Phe	Gly	Met	Ser	Gly	Thr	Asn	Cys	His	Leu	Val	Val
465					470					475					480
Ala	Glu	Ala	Pro	Val	Arg	Gln	Asn	Glu	Lys	Ser	Glu	Asn	Ala	Pro	Glu
				485					490					495	
Arg	Pro	Cys	His	Ile	Leu	Thr	Leu	Ser	Ala	Lys	Thr	Glu	Ala	Ala	Leu
		500						505						510	
Asn	Ala	Leu	Val	Ala	Arg	Tyr	Met	Ala	Phe	Leu	Arg	Glu	Ala	Pro	Ala
	515						520							525	
Ile	Ser	Leu	Ala	Asp	Leu	Cys	Tyr	Ser	Ala	Asn	Val	Gly	Arg	Asn	Leu
	530					535					540				
Phe	Ala	His	Arg	Leu	Ser	Phe	Ile	Ser	Glu	Asn	Ile	Ala	Gln	Leu	Ser
545					550					555					560
Glu	Gln	Leu	Glu	His	Cys	Pro	Gln	Gln	Ala	Thr	Met	Pro	Thr	Gln	His
				565					570					575	
Asn	Val	Ile	Leu	Asp	Asn	Gln	Leu	Ser	Pro	Gln	Ile	Ala	Phe	Leu	Phe
		580						585						590	
Thr	Gly	Gln	Gly	Ser	Gln	Tyr	Ile	Asn	Met	Gly	Arg	Glu	Leu	Tyr	Glu
	595						600							605	
Thr	Gln	Pro	Thr	Phe	Arg	Arg	Ile	Met	Asp	Glu	Cys	Asp	Asp	Ile	Leu
	610					615					620				
His	Pro	Leu	Leu	Gly	Glu	Ser	Ile	Leu	Asn	Ile	Leu	Tyr	Thr	Ser	Pro
625					630					635					640
Ser	Lys	Leu	Asn	Gln	Thr	Val	Tyr	Thr	Gln	Pro	Ala	Leu	Phe	Ala	Phe
				645					650					655	
Glu	Tyr	Ala	Leu	Ala	Lys	Leu	Trp	Ile	Ser	Trp	Gly	Ile	Glu	Pro	Asp
		660						665						670	
Val	Val	Leu	Gly	His	Ser	Val	Gly	Glu	Tyr	Val	Ala	Ala	Cys	Leu	Ala
	675						680						685		
Gly	Val	Phe	Ser	Leu	Glu	Asp	Gly	Leu	Lys	Leu	Ile	Ala	Ser	Arg	Gly
	690					695					700				
Cys	Leu	Met	Gln	Ala	Leu	Pro	Pro	Gly	Lys	Met	Leu	Ser	Ile	Arg	Ser
705					710					715					720
Asn	Glu	Ile	Gly	Val	Lys	Ala	Leu	Ile	Ala	Pro	Tyr	Ser	Ala	Glu	Val
				725					730					735	
Ser	Ile	Ala	Ala	Ile	Asn	Gly	Gln	Gln	Ser	Val	Val	Ile	Ser	Gly	Lys
		740						745						750	
Ala	Glu	Ile	Ile	Asp	Asn	Leu	Ala	Ala	Glu	Phe	Ala	Ser	Glu	Gly	Ile
	755						760						765		
Lys	Thr	His	Leu	Ile	Thr	Val	Ser	His	Ala	Phe	His	Ser	Pro	Met	Met
	770					775					780				
Thr	Pro	Met	Leu	Lys	Ala	Phe	Arg	Asp	Val	Ala	Ser	Thr	Ile	Ser	Tyr
785					790					795					800
Arg	Ser	Pro	Ser	Leu	Ser	Leu	Ile	Ser	Asn	Gly	Thr	Gly	Gln	Leu	Ala

					805					810					815
Thr	Lys	Glu	Val	Ala	Thr	Pro	Asp	Tyr	Trp	Val	Arg	His	Val	His	Ser
			820					825					830		
Thr	Val	Arg	Phe	Ala	Asp	Gly	Ile	Ala	Thr	Leu	Ala	Glu	Gln	Asn	Thr
			835					840					845		
Asp	Ile	Leu	Leu	Glu	Val	Gly	Pro	Lys	Pro	Ile	Leu	Leu	Gly	Met	Ala
			850					855					860		
Lys	Gln	Ile	Tyr	Ser	Glu	Asn	Gly	Ser	Ala	Ser	His	Pro	Leu	Met	Leu
865						870					875				880
Pro	Ser	Leu	Arg	Glu	Asp	Gly	Asn	Asp	Trp	Gln	Gln	Met	Leu	Ser	Thr
					885					890					895
Cys	Gly	Gln	Leu	Val	Val	Asn	Gly	Val	Lys	Ile	Asp	Trp	Ala	Gly	Phe
			900					905					910		
Asp	Lys	Asp	Tyr	Ser	Arg	His	Lys	Ile	Leu	Leu	Pro	Thr	Tyr	Pro	Phe
			915					920					925		
Gln	Arg	Glu	Arg	Tyr	Trp	Ile	Glu	Ser	Ser	Val	Lys	Lys	Pro	Gln	Lys
			930					935					940		
Gln	Glu	Leu	Arg	Pro	Met	Leu	Asp	Lys	Met	Ile	Arg	Leu	Pro	Ser	Glu
945						950					955				960
Asn	Lys	Val	Val	Phe	Glu	Thr	Glu	Phe	Gly	Val	Arg	Gln	Met	Pro	His
					965					970					975
Ile	Ser	Asp	His	Gln	Ile	Tyr	Gly	Glu	Val	Ile	Val	Pro	Gly	Ala	Val
			980						985					990	
Leu	Ala	Ser	Leu	Ile	Phe	Asn	Ala	Ala	Gln	Val	Leu	Tyr	Pro	Asp	Tyr
			995					1000					1005		
Gln	His	Glu	Leu	Thr	Asp	Ile	Ala	Phe	Tyr	Gln	Pro	Ile	Ile	Phe	
			1010					1015				1020			
His	Asp	Asp	Asp	Thr	Val	Ile	Val	Gln	Ala	Ile	Phe	Ser	Pro	Asp	
			1025					1030				1035			
Lys	Ser	Gln	Glu	Asn	Gln	Ser	His	Gln	Thr	Phe	Pro	Pro	Met	Ser	
			1040					1045				1050			
Phe	Gln	Ile	Ile	Ser	Phe	Met	Pro	Asp	Gly	Pro	Leu	Glu	Asn	Lys	
			1055					1060				1065			
Pro	Lys	Val	His	Val	Thr	Gly	Cys	Leu	Arg	Met	Leu	Arg	Asp	Ala	
			1070					1075				1080			
Gln	Pro	Pro	Thr	Leu	Ser	Pro	Thr	Glu	Ile	Arg	Gln	Arg	Cys	Pro	
			1085					1090				1095			
His	Thr	Val	Asn	Gly	His	Asp	Trp	Tyr	Asn	Ser	Leu	Val	Lys	Gln	
			1100					1105				1110			
Lys	Phe	Glu	Met	Gly	Pro	Ser	Phe	Arg	Trp	Val	Gln	Gln	Leu	Trp	
			1115					1120				1125			
His	Gly	Glu	Asn	Glu	Ala	Leu	Thr	Arg	Leu	His	Ile	Pro	Asp	Val	
			1130					1135				1140			
Val	Gly	Ser	Val	Ser	Gly	His	Gln	Leu	His	Gly	Ile	Leu	Leu		

Ile	Glu	Trp	Gln	Ala	Gln	Ser	Thr	Ser	Leu	Val	Pro	Ser	Asp	Gly
1250						1255					1260			
Thr	Ile	Ser	Gly	Ser	Cys	Leu	Val	Leu	Ser	Asp	Gln	His	Gly	Thr
1265						1270					1275			
Gly	Ala	Ala	Leu	Ala	Gln	Arg	Leu	Asp	Asn	Ala	Gly	Val	Pro	Val
1280						1285					1290			
Thr	Met	Ile	Tyr	Ala	Asp	Leu	Ile	Leu	Asp	Asn	Tyr	Glu	Leu	Ile
1295						1300					1305			
Phe	Arg	Thr	Leu	Pro	Asp	Leu	Gln	Gln	Val	Val	Tyr	Leu	Trp	Gly
1310						1315					1320			
Leu	Asp	Gln	Lys	Glu	Asp	Cys	His	Pro	Met	Lys	Gln	Ala	Glu	Asp
1325						1330					1335			
Asn	Cys	Thr	Ser	Val	Leu	Tyr	Leu	Val	Gln	Ala	Leu	Leu	Asn	Thr
1340						1345					1350			
Tyr	Ser	Thr	Pro	Pro	Ser	Leu	Leu	Ile	Val	Thr	Cys	Asp	Ala	Gln
1355						1360					1365			
Ala	Val	Val	Glu	Gln	Asp	Arg	Val	Asn	Gly	Phe	Ala	Gln	Ser	Ser
1370						1375					1380			
Leu	Leu	Gly	Leu	Ala	Lys	Val	Ile	Met	Leu	Glu	His	Pro	Glu	Leu
1385						1390					1395			
Ser	Cys	Val	Tyr	Met	Asp	Val	Glu	Ala	Gly	Tyr	Leu	Gln	Gln	Asp
1400						1405					1410			
Val	Ala	Asn	Thr	Ile	Phe	Thr	Gln	Leu	Lys	Arg	Gly	His	Leu	Ser
1415						1420					1425			
Lys	Asp	Gly	Glu	Glu	Ser	Gln	Leu	Ala	Trp	Arg	Asn	Gly	Gln	Ala
1430						1435					1440			
Tyr	Val	Ala	Arg	Leu	Ser	Gln	Tyr	Lys	Pro	Lys	Ser	Glu	Gln	Leu
1445						1450					1455			
Val	Glu	Ile	Arg	Ser	Asp	Arg	Ser	Tyr	Leu	Ile	Thr	Gly	Gly	Arg
1460						1465					1470			
Gly	Gly	Val	Gly	Leu	Gln	Ile	Ala	Arg	Trp	Leu	Val	Glu	Lys	Gly
1475						1480					1485			
Ala	Lys	His	Leu	Val	Leu	Leu	Gly	Arg	Ser	Gln	Thr	Ser	Ser	Glu
1490						1495					1500			
Val	Ser	Leu	Val	Leu	Asp	Glu	Leu	Glu	Ser	Ala	Gly	Ala	Gln	Ile
1505						1510					1515			
Ile	Val	Ala	Gln	Ala	Asp	Ile	Ser	Asp	Glu	Lys	Val	Leu	Ala	Gln
1520						1525					1530			
Ile	Leu	Thr	Asn	Leu	Thr	Val	Pro	Leu	Cys	Gly	Val	Ile	His	Ala
1535						1540					1545			
Ala	Gly	Val	Leu	Asp	Asp	Ala	Ser	Leu	Leu	Gln	Gln	Thr	Pro	Ala
1550						1555					1560			
Lys	Leu	Lys	Lys	Val	Leu	Leu	Pro	Lys	Ala	Glu	Gly	Ala	Trp	Ile
1565						1570					1575			
Leu	His	Asn	Leu	Thr	Leu	Glu	Gln	Arg	Leu	Asp	Phe	Phe	Val	Leu
1580						1585					1590			
Phe	Ser	Ser	Ala	Ser	Ser	Leu	Leu	Gly	Ala	Pro	Gly	Gln	Ala	Asn
1595						1600					1605			
Tyr	Ser	Ala	Ala	Asn	Ala	Phe	Leu	Asp	Gly	Leu	Ala	Ala	Tyr	Arg
1610						1615					1620			
Arg	Gly	Arg	Gly	Leu	Pro	Cys	Leu	Ser	Ile	Cys	Trp	Gly	Ala	Trp
1625						1630					1635			
Asp	Gln	Val	Gly	Met	Ala	Ala	Arg	Gln	Gly	Leu	Leu	Asp	Lys	Leu
1640						1645					1650			
Pro	Gln	Arg	Gly	Glu	Glu	Ala	Ile	Pro	Leu	Gln	Lys	Gly	Leu	Asp
1655						1660					1665			
Leu	Phe	Gly	Glu	Leu	Leu	Asn	Glu	Pro	Ala	Ala	Gln	Ile	Gly	Val

1670	Ile Pro	Ile Gln Trp Thr	1675	Arg Phe Leu Asp His	1680	Gln Lys Gly Asn
1685	Leu Pro	Phe Tyr Glu Lys	1690	Phe Ser Lys Ser Ser	1695	Arg Lys Ala Gln
1700	Ser Tyr	Asp Ser Met Ala	1705	Ser His Thr Glu	1710	Asp Ile Gln Arg
1715	Lys Leu	Lys Gln Ala Ala	1720	Gln Asp Arg Pro	1725	Lys Leu Leu Glu
1730	Val His	Leu Arg Ser Gln	1735	Val Ala Gln Leu Leu	1740	Gly Ile Asn Val
1745	Ala Glu	Leu Pro Asn Glu	1750	Gly Ile Gly Phe	1755	Val Thr Leu Gly
1760	Leu Asp	Ser Leu Thr Ser	1765	Ile Glu Leu Arg Asn	1770	Ser Leu Gln Arg
1775	Thr Leu	Asp Cys Ser Leu	1780	Pro Val Thr Phe Ala	1785	Phe Asp Tyr Pro
1790	Thr Ile	Glu Ile Ala Val	1795	Lys Tyr Leu Thr Gln	1800	Val Val Ile Ala
1805	Pro Met	Glu Ser Thr Ala	1810	Ser Gln Gln Thr Asp	1815	Ser Leu Ser Ala
1820	Met Phe	Thr Asp Thr Ser	1825	Ser Ile Gly Arg Ile	1830	Leu Asp Asn Glu
1835	Thr Asp	Val Leu Asp Ser	1840	Glu Met Gln Ser Asp	1845	Glu Asp Glu Ser
1850	Leu Ser	Thr Leu Ile Gln	1855	Lys Leu Ser Thr His	1860	Leu Asp
1865			1870		1875	

<210> 83

<211> 4074

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 83

atgaacgctt	tgtcagaaaa	tcaggtaact	tctatagtc	agaaggcatt	gaacaaaata	60
gaggagttac	aagccgaact	tgaccgttta	aaatacgcgc	aacgggaacc	aatcgccatc	120
attggaatgg	gctgtcgctt	tcctgggtgca	gacacacctg	aagctttttg	gaaattattg	180
cacaatgggg	ttgatgctat	ccaagagatt	ccaaaaagcc	gttgggatat	tgacgactat	240
tatgatccca	caccagcaac	acccggcaaa	atgtatacac	gttttggtgg	ttttctcgac	300
caaatagcag	ccttcgaccc	tgagttcttt	cgcatttcta	ctcgtgaggc	aatcagctta	360
gaccctcaac	agagattgct	tctggaagtg	agttgggaag	ccttagaacg	ggctgggctg	420
acaggcaata	aactgactac	acaaacaggt	gtctttgttg	gcatcagtga	aagtgattat	480
cgtgatttga	ttatgcgtaa	tggttctgac	ctagatgtat	attctgggtc	aggtaactgc	540
catagtacag	ccagcggggc	tttatcttat	tatttgggac	ttactggacc	caatttgtcc	600
cttgataccg	cctgttcgtc	ctctttgggt	tgtgtggcat	tggctgtcaa	gagcctacgt	660
caacaggagt	gtgatttggc	attggcgggt	ggtgtacaga	tacaagtgat	accagatggc	720
tttatcaaag	cctgtcaatc	ccgtatgttg	tcgcctgatg	gacggtgcaa	aacatttgat	780
ttccaggcag	atggttatgc	ccgtgctgag	gggtgtggga	tggtagttct	caaacgccta	840
tccgatgcaa	ttgctgacaa	tgataatatc	ctggccttga	ttcgtgggtg	cgcagtcaat	900
catgatggct	acacgagtgg	attaaccgtt	cccagtggtc	cctcacaacg	ggcggtgatc	960
caacaggcat	tagcggatgc	tggaatacac	ccggatcaaa	ttagctatat	tgaggcacat	1020
ggcacaggta	catccttagg	cgatcctatt	gaaatgggtg	cgattgggca	agtctttggt	1080
caacgctcac	agatgctttt	cgtcgggttc	gtcaagacga	atattggtca	tactgaggct	1140
gctgctggta	ttgctgggtc	catcaagggt	gtactctcaa	tgacgacg	tgaaatcca	1200
gcaaacttac	acttcgacca	gccaaatcct	tatattaact	gggatcaatt	accagtcagt	1260
atcccaacag	aaacaatacc	ttggtctact	agcgatcgct	ttgcaggagt	cagtagcttt	1320
ggcttttagtg	gcacaaactc	tcatatcgta	ctagaggcag	ccccaaacat	agagcaacct	1380
actgatgata	ttaatcaaac	gccgcatatt	ttgaccttag	ctgcaaaaac	acccgcagcc	1440

10

```

ctgcaagaac tggctcggcg ttatgcgact cagatagaga cctctcccga tgttcctctg 1500
gcggaacatt gtttcacagc acacataggg cgtaaacatt ttaaacaatag gtttgcggtg 1560
gtcacggaat ctaaagagca actgcggttg caattggatg catttgacaca atcagggggg 1620
gtggggcgag aagtcaaate gctaccaaag atagccttcc tttttacagg tcaagggtca 1680
cagtatgtgg gaatgggtcg tcaactttac gaaaaccaac ctaccttccg aaaagcactc 1740
gcccattgtg atgacatcct gcgtgctggt gcatatttcg accgatcact actttcgatt 1800
ctctaccagc agggaaaatc agaagccatt caccaaaccg cttatactca gcccgcgctt 1860
tttgctcttg agtatgcat cgctcagttg tggcactcct ggggtatcaa accagatatc 1920
gtgatggggc atagtgtagg tgaatacgtc gccgcttggt tggcgggcat attttcttta 1980
gaggatgggc tgaactaat tgctactcgt ggtcgtctga tgcaatccct acctcaagac 2040
ggaacgatgg tttcttcttt ggcaagtga gctcgtatcc aggaagctat tacaccttac 2100
cgagatgatg tgtcaatcgc agcgataaat gggacagaaa gcgtgggttat ctctggcaaa 2160
cgacactctg tgatggcaat tgctgaacaa ctcgccaccg ttggcatcaa gacacgcaa 2220
ctgacgggtt cccatgcctt ccattcacca cttatgacac ccattcttga tgagtccgc 2280
caggtggcag ccagtatcac ctatcaccag cccaagttgc tacttgtctc caacgtctcc 2340
gggaaagtgg ccggccctga aatcaccaga ccagattact gggtacgcca tgtccgtgag 2400
gcagtgcgct ttgccgatgg agtgaggacg ctgaatgaac aaggtgtcaa tatctttctg 2460
gaaatcggtt ctaccgctac cctgttgggc atggcactgc gagtaaatga ggaagattca 2520
aatgcctcaa aaggaacttc gtcttgctac ctgccagtt tacgggaaag ccagaaggat 2580
tgtcagcaga tgttactatg tctgggtgag ttgtacgtac atggatatga tattgattgg 2640
ggtgcattta atcggggata tcaaggacgc aagtgatat tgccaacctc tccgtttcag 2700
cgacaacggt attggcttcc cgaccctaag ttggcacaaa gttccgattt agataccttt 2760
caagctcaga gcagcgcac atcacaaaat cctagcgtg tgccacttt actgatggaa 2820
tatttgcaag caggtgatgt ccaatcttta gttgggcttt tggatgatga acggaaactc 2880
tctgtgctg aacgaattgc actaccaggt attttggagt ttttggtaga ggaacaacag 2940
cgacaataaa gctcaaccac aactcctcaa acagttttac aaaaaataag tcaaaacttc 3000
catgaggaca gatatgaaat attgaagaac ctgatcaaat ctgaaatcga aacgattatc 3060
aaaagtgttc cctccgatga acaaatgttt tctgacttag gaattgattc cttgatggcg 3120
atcgaaactg gtaataagct ccgttctgct ataggggttg aactgccagt ggcaatagta 3180
tttgaccatc ccacgattaa gcagttaact aacttcgtac tggacagaat tgtgccgcag 3240
gcagaccaaa aggacgttcc caccgaatcc ttgttgctt ctaaacagga gatatcagtt 3300
gaggagcagt cttttgcaat taccaagctg ggcttatccc ctgcttccca ctccctgcat 3360
cttctcccat ggacggttag acctgcggtg atggcagatg taacaaaact aagccaactt 3420
gaaagagagg cctatggctg gatcgagaaa ggagcgatcg cccgcgccca tctcattgcc 3480
gatcgcatca atttactcaa cagtgggtgat atgccttggg tctgggtaat ggagcgatca 3540
ggagagttgg gcgctgggca ggtgctacaa ccgacatctg ttgatccata tacttatgga 3600
agttgggatg aagtaactga ccaaggtaaa ctgcaagcaa ccttcgaccc aagtggacgc 3660
aatgtgtata ttgtcgcggt tgggtctagc aacctccca cggtagccag ccacctcatg 3720
acgcttcaga ctttattgat gctgcgggaa actggtcgtg acacaatctt tgtctgtctg 3780
gcaatgccag gttatgcaa ataccacagt caaacaggaa aatcgccgga agagtatatt 3840
gcgctgactg acgaggatgg tatcccaatg gacgagttta ttgcactttc tgtctacgac 3900
tggectgtta ccccatcggt tcgtgttctg cgagacggtt atccacctga tcgagattct 3960
ggtggtcacg cagttagtac ggttttccag ctcaatgatt tcgatggagc gatcgaagaa 4020
acatatcgtc gtattatccg ccattgccgat gtccttggtc tcgaaagagg ctaa 4074

```

<210> 84

<211> 1357

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 84

```

Met Asn Ala Leu Ser Glu Asn Gln Val Thr Ser Ile Val Lys Lys Ala
1      5      10      15
Leu Asn Lys Ile Glu Glu Leu Gln Ala Glu Leu Asp Arg Leu Lys Tyr
20     25     30
Ala Gln Arg Glu Pro Ile Ala Ile Ile Gly Met Gly Cys Arg Phe Pro
35     40     45
10 Gly Ala Asp Thr Pro Glu Ala Phe Trp Lys Leu Leu His Asn Gly Val

```

50	55	60
Asp Ala Ile Gln Glu Ile	Pro Lys Ser Arg Trp	Asp Ile Asp Asp Tyr
65	70	75
Tyr Asp Pro Thr Pro	Ala Thr Pro Gly Lys	Met Tyr Thr Arg Phe Gly
85	90	95
Gly Phe Leu Asp Gln Ile	Ala Ala Phe Asp Pro	Glu Phe Phe Arg Ile
100	105	110
Ser Thr Arg Glu Ala Ile	Ser Leu Asp Pro	Gln Gln Arg Leu Leu Leu
115	120	125
Glu Val Ser Trp Glu Ala	Leu Glu Arg Ala Gly	Leu Thr Gly Asn Lys
130	135	140
Leu Thr Thr Gln Thr Gly	Val Phe Val Gly Ile	Ser Glu Ser Asp Tyr
145	150	155
Arg Asp Leu Ile Met Arg	Asn Gly Ser Asp Leu	Asp Val Tyr Ser Gly
165	170	175
Ser Gly Asn Cys His Ser	Thr Ala Ser Gly Arg	Leu Ser Tyr Tyr Leu
180	185	190
Gly Leu Thr Gly Pro Asn	Leu Ser Leu Asp Thr	Ala Cys Ser Ser Ser
195	200	205
Leu Val Cys Val Ala Leu	Ala Val Lys Ser Leu	Arg Gln Gln Glu Cys
210	215	220
Asp Leu Ala Leu Ala Gly	Gly Val Gln Ile Gln	Val Ile Pro Asp Gly
225	230	235
Phe Ile Lys Ala Cys Gln	Ser Arg Met Leu Ser	Pro Asp Gly Arg Cys
245	250	255
Lys Thr Phe Asp Phe Gln	Ala Asp Gly Tyr Ala	Arg Ala Glu Gly Cys
260	265	270
Gly Met Val Val Leu Lys	Arg Leu Ser Asp Ala	Ile Ala Asp Asn Asp
275	280	285
Asn Ile Leu Ala Leu Ile	Arg Gly Ala Ala Val	Asn His Asp Gly Tyr
290	295	300
Thr Ser Gly Leu Thr Val	Pro Ser Gly Pro Ser	Gln Arg Ala Val Ile
305	310	315
Gln Gln Ala Leu Ala Asp	Ala Gly Ile His Pro	Asp Gln Ile Ser Tyr
325	330	335
Ile Glu Ala His Gly Thr	Gly Thr Ser Leu Gly	Asp Pro Ile Glu Met
340	345	350
Gly Ala Ile Gly Gln Val	Phe Gly Gln Arg Ser	Gln Met Leu Phe Val
355	360	365
Gly Ser Val Lys Thr Asn	Ile Gly His Thr Glu	Ala Ala Ala Gly Ile
370	375	380
Ala Gly Leu Ile Lys Val	Val Leu Ser Met Gln	His Gly Glu Ile Pro
385	390	395
Ala Asn Leu His Phe Asp	Gln Pro Ser Pro Tyr	Ile Asn Trp Asp Gln
405	410	415
Leu Pro Val Ser Ile Pro	Thr Glu Thr Ile Pro	Trp Ser Thr Ser Asp
420	425	430
Arg Phe Ala Gly Val Ser	Ser Phe Gly Phe Ser	Gly Thr Asn Ser His
435	440	445
Ile Val Leu Glu Ala Ala	Pro Asn Ile Glu Gln	Pro Thr Asp Asp Ile
450	455	460
Asn Gln Thr Pro His Ile	Leu Thr Leu Ala Ala	Lys Thr Pro Ala Ala
465	470	475
Leu Gln Glu Leu Ala Arg	Arg Tyr Ala Thr Gln	Ile Glu Thr Ser Pro
485	490	495
Asp Val Pro Leu Ala Asp	Ile Cys Phe Thr Ala	His Ile Gly Arg Lys
500	505	510

His	Phe	Lys	His	Arg	Phe	Ala	Val	Val	Thr	Glu	Ser	Lys	Glu	Gln	Leu
		515					520					525			
Arg	Leu	Gln	Leu	Asp	Ala	Phe	Ala	Gln	Ser	Gly	Gly	Val	Gly	Arg	Glu
	530				535						540				
Val	Lys	Ser	Leu	Pro	Lys	Ile	Ala	Phe	Leu	Phe	Thr	Gly	Gln	Gly	Ser
545					550					555					560
Gln	Tyr	Val	Gly	Met	Gly	Arg	Gln	Leu	Tyr	Glu	Asn	Gln	Pro	Thr	Phe
				565					570					575	
Arg	Lys	Ala	Leu	Ala	His	Cys	Asp	Asp	Ile	Leu	Arg	Ala	Gly	Ala	Tyr
			580					585					590		
Phe	Asp	Arg	Ser	Leu	Leu	Ser	Ile	Leu	Tyr	Pro	Glu	Gly	Lys	Ser	Glu
		595					600					605			
Ala	Ile	His	Gln	Thr	Ala	Tyr	Thr	Gln	Pro	Ala	Leu	Phe	Ala	Leu	Glu
	610					615					620				
Tyr	Ala	Ile	Ala	Gln	Leu	Trp	His	Ser	Trp	Gly	Ile	Lys	Pro	Asp	Ile
625					630					635					640
Val	Met	Gly	His	Ser	Val	Gly	Glu	Tyr	Val	Ala	Ala	Cys	Val	Ala	Gly
				645					650					655	
Ile	Phe	Ser	Leu	Glu	Asp	Gly	Leu	Lys	Leu	Ile	Ala	Thr	Arg	Gly	Arg
			660				665						670		
Leu	Met	Gln	Ser	Leu	Pro	Gln	Asp	Gly	Thr	Met	Val	Ser	Ser	Leu	Ala
		675					680					685			
Ser	Glu	Ala	Arg	Ile	Gln	Glu	Ala	Ile	Thr	Pro	Tyr	Arg	Asp	Asp	Val
	690					695					700				
Ser	Ile	Ala	Ala	Ile	Asn	Gly	Thr	Glu	Ser	Val	Val	Ile	Ser	Gly	Lys
705					710					715					720
Arg	Thr	Ser	Val	Met	Ala	Ile	Ala	Glu	Gln	Leu	Ala	Thr	Val	Gly	Ile
			725						730					735	
Lys	Thr	Arg	Gln	Leu	Thr	Val	Ser	His	Ala	Phe	His	Ser	Pro	Leu	Met
			740					745					750		
Thr	Pro	Ile	Leu	Asp	Glu	Phe	Arg	Gln	Val	Ala	Ala	Ser	Ile	Thr	Tyr
		755					760					765			
His	Gln	Pro	Lys	Leu	Leu	Leu	Val	Ser	Asn	Val	Ser	Gly	Lys	Val	Ala
	770					775					780				
Gly	Pro	Glu	Ile	Thr	Arg	Pro	Asp	Tyr	Trp	Val	Arg	His	Val	Arg	Glu
785					790					795					800
Ala	Val	Arg	Phe	Ala	Asp	Gly	Val	Arg	Thr	Leu	Asn	Glu	Gln	Gly	Val
			805						810					815	
Asn	Ile	Phe	Leu	Glu	Ile	Gly	Ser	Thr	Ala	Thr	Leu	Leu	Gly	Met	Ala
			820					825					830		
Leu	Arg	Val	Asn	Glu	Glu	Asp	Ser	Asn	Ala	Ser	Lys	Gly	Thr	Ser	Ser
		835					840					845			
Cys	Tyr	Leu	Pro	Ser	Leu	Arg	Glu	Ser	Gln	Lys	Asp	Cys	Gln	Gln	Met
	850					855					860				
Phe	Thr	Ser	Leu	Gly	Glu	Leu	Tyr	Val	His	Gly	Tyr	Asp	Ile	Asp	Trp
865					870					875					880
Gly	Ala	Phe	Asn	Arg	Gly	Tyr	Gln	Gly	Arg	Lys	Val	Ile	Leu	Pro	Thr
			885						890					895	
Tyr	Pro	Phe	Gln	Arg	Gln	Arg	Tyr	Trp	Leu	Pro	Asp	Pro	Lys	Leu	Ala
			900					905					910		
Gln	Ser	Ser	Asp	Leu	Asp	Thr	Phe	Gln	Ala	Gln	Ser	Ser	Ala	Ser	Ser
		915					920					925			
Gln	Asn	Pro	Ser	Ala	Val	Ser	Thr	Leu	Leu	Met	Glu	Tyr	Leu	Gln	Ala
	930					935					940				
Gly	Asp	Val	Gln	Ser	Leu	Val	Gly	Leu	Leu	Asp	Asp	Glu	Arg	Lys	Leu
945					950					955					960
Ser	Ala	Ala	Glu	Arg	Ile	Ala	Leu	Pro	Ser	Ile	Leu	Glu	Phe	Leu	Val

965										970					975				
Glu	Glu	Gln	Gln	Arg	Gln	Ile	Ser	Ser	Thr	Thr	Thr	Thr	Pro	Gln	Thr	Val			
980										985					990				
Leu	Gln	Lys	Ile	Ser	Gln	Thr	Ser	His	Glu	Asp	Arg	Tyr	Glu	Ile	Leu				
995										1000					1005				
Lys	Asn	Leu	Ile	Lys	Ser	Glu	Ile	Glu	Thr	Ile	Ile	Lys	Ser	Val					
	1010					1015					1020								
Pro	Ser	Asp	Glu	Gln	Met	Phe	Ser	Asp	Leu	Gly	Ile	Asp	Ser	Leu					
	1025					1030					1035								
Met	Ala	Ile	Glu	Leu	Arg	Asn	Lys	Leu	Arg	Ser	Ala	Ile	Gly	Leu					
	1040					1045					1050								
Glu	Leu	Pro	Val	Ala	Ile	Val	Phe	Asp	His	Pro	Thr	Ile	Lys	Gln					
	1055					1060					1065								
Leu	Thr	Asn	Phe	Val	Leu	Asp	Arg	Ile	Val	Pro	Gln	Ala	Asp	Gln					
	1070					1075					1080								
Lys	Asp	Val	Pro	Thr	Glu	Ser	Leu	Phe	Ala	Ser	Lys	Gln	Glu	Ile					
	1085					1090					1095								
Ser	Val	Glu	Glu	Gln	Ser	Phe	Ala	Ile	Thr	Lys	Leu	Gly	Leu	Ser					
	1100					1105					1110								
Pro	Ala	Ser	His	Ser	Leu	His	Leu	Pro	Pro	Trp	Thr	Val	Arg	Pro					
	1115					1120					1125								
Ala	Val	Met	Ala	Asp	Val	Thr	Lys	Leu	Ser	Gln	Leu	Glu	Arg	Glu					
	1130					1135					1140								
Ala	Tyr	Gly	Trp	Ile	Gly	Glu	Gly	Ala	Ile	Ala	Pro	Pro	His	Leu					
	1145					1150					1155								
Ile	Ala	Asp	Arg	Ile	Asn	Leu	Leu	Asn	Ser	Gly	Asp	Met	Pro	Trp					
	1160					1165					1170								
Phe	Trp	Val	Met	Glu	Arg	Ser	Gly	Glu	Leu	Gly	Ala	Trp	Gln	Val					
	1175					1180					1185								
Leu	Gln	Pro	Thr	Ser	Val	Asp	Pro	Tyr	Thr	Tyr	Gly	Ser	Trp	Asp					
	1190					1195					1200								
Glu	Val	Thr	Asp	Gln	Gly	Lys	Leu	Gln	Ala	Thr	Phe	Asp	Pro	Ser					
	1205					1210					1215								
Gly	Arg	Asn	Val	Tyr	Ile	Val	Ala	Gly	Gly	Ser	Ser	Asn	Leu	Pro					
	1220					1225					1230								
Thr	Val	Ala	Ser	His	Leu	Met	Thr	Leu	Gln	Thr	Leu	Leu	Met	Leu					
	1235					1240					1245								
Arg	Glu	Thr	Gly	Arg	Asp	Thr	Ile	Phe	Val	Cys	Leu	Ala	Met	Pro					
	1250					1255					1260								
Gly	Tyr	Ala	Lys	Tyr	His	Ser	Gln	Thr	Gly	Lys	Ser	Pro	Glu	Glu					
	1265					1270					1275								
Tyr	Ile	Ala	Leu	Thr	Asp	Glu	Asp	Gly	Ile</										

<210> 85

<211> 1437

5 <212> ADN

<212> ADN

<213> cylindrospermopsis raciborskii awt205

<400> 85

```

atgaataaaa aacaggtaga cacattgtta atacacgctc atctttttac catgcagggc 60
aatggcctgg gatataattgc cgatggggca attgcggttc agggtagcca gatcgtagca 120
gtggattcga cagaggcttt gctgagtcac tttgaaggaa ataaaacaat taatgcggta 180
aattgtgcag tgttgccctgg actaattgat gctcatatac atacgacttg tgctattctg 240
cgtggagtgg cacaggatgt aaccaattgg ctaatggacg cgacaattcc ttatgcactt 300
cagatgacac ccgcagtaaa tatagccgga acgcgcttga gtgtactcga agggctgaaa 360
gcaggaacaa ccacattcgg cgattctgag actccttacc cgctctgggg agagtttttc 420
gatgaaattg gggtagctgc tattctatcc cctgccttta acgcctttcc actagaatgg 480
tcggcatgga aggagggaga cctctatccc ttcatatga aggcaggacg acgtgggtatg 540
gaagaggctg tggattttgc ttgtgcatgg aatggagccg cagagggacg tatcaccact 600
atgttgggac tacaggcggc ggatatgcta ccactggaga tcctacacgc agctaaagag 660
attgcccac ccggaaggcctt aatgctgcat attcatgtgg cccagggaga tcgagaaaca 720
aaacaaattg tcaaacgata tggtaagcgt ccgatcgcat ttctagctga aattggctac 780
ttggacgaac agttgctggc agttcacctc accgatgcca cagatgaaga agtgatacaa 840
gtagccaaaa gtggtgctgg catggcactc tgttcggggc ctattggcat cattgacggt 900
cttgttccgc ccgctcatgt ttttcgacaa gcaggcgggt ccgttgact cggttctgat 960
caagcctgtg gcaacaactg ttgtaacatc ttcaatgaaa tgaagctgac cgccttattc 1020
aacaaaaata aatatcatga tccaaccatt atgccggctt ggaagtcct gcgtatggct 1080
accatcgaag gagcgcaggc gattggttta gatcacaaga ttggctctct tcaagtgggc 1140
aaagaagccg acctgatctt aatagacctc agttccccta acctctgcc caccctgctc 1200
aaccctattc gtaaccttgt acctaactg gtgtatgctg cttcaggaca tgaagttaaa 1260
agcgtcatgg tggcgggaaa acttttagtg gaagactacc aagtcctcac ggtagatgag 1320
tccgctattc tcgctgaagc gcaagtacaa gctcaacaac tctgccaacg tgtgaccgct 1380
gaccccatcc acaaaaagat ggtgttaatg gaagcgatgg ctaagggtaa attatag 1437

```

<210> 86

<211> 478

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 86

```

Met Asn Lys Lys Gln Val Asp Thr Leu Leu Ile His Ala His Leu Phe
1 5 10 15
Thr Met Gln Gly Asn Gly Leu Gly Tyr Ile Ala Asp Gly Ala Ile Ala
20 25 30
Val Gln Gly Ser Gln Ile Val Ala Val Asp Ser Thr Glu Ala Leu Leu
35 40 45
Ser His Phe Glu Gly Asn Lys Thr Ile Asn Ala Val Asn Cys Ala Val
50 55 60
Leu Pro Gly Leu Ile Asp Ala His Ile His Thr Thr Cys Ala Ile Leu
65 70 75 80
Arg Gly Val Ala Gln Asp Val Thr Asn Trp Leu Met Asp Ala Thr Ile
85 90 95
Pro Tyr Ala Leu Gln Met Thr Pro Ala Val Asn Ile Ala Gly Thr Arg
100 105 110
Leu Ser Val Leu Glu Gly Leu Lys Ala Gly Thr Thr Thr Phe Gly Asp
115 120 125
Ser Glu Thr Pro Tyr Pro Leu Trp Gly Glu Phe Phe Asp Glu Ile Gly
130 135 140
Val Arg Ala Ile Leu Ser Pro Ala Phe Asn Ala Phe Pro Leu Glu Trp
145 150 155 160
Ser Ala Trp Lys Glu Gly Asp Leu Tyr Pro Phe Asp Met Lys Ala Gly
165 170 175
Arg Arg Gly Met Glu Glu Ala Val Asp Phe Ala Cys Ala Trp Asn Gly
180 185 190
10 Ala Ala Glu Gly Arg Ile Thr Thr Met Leu Gly Leu Gln Ala Ala Asp

```

```

      195              200              205
Met Leu Pro Leu Glu Ile Leu His Ala Ala Lys Glu Ile Ala Gln Arg
  210              215              220
Glu Gly Leu Met Leu His Ile His Val Ala Gln Gly Asp Arg Glu Thr
225              230              235              240
Lys Gln Ile Val Lys Arg Tyr Gly Lys Arg Pro Ile Ala Phe Leu Ala
      245              250              255
Glu Ile Gly Tyr Leu Asp Glu Gln Leu Leu Ala Val His Leu Thr Asp
      260              265              270
Ala Thr Asp Glu Glu Val Ile Gln Val Ala Lys Ser Gly Ala Gly Met
      275              280              285
Ala Leu Cys Ser Gly Ala Ile Gly Ile Ile Asp Gly Leu Val Pro Pro
      290              295              300
Ala His Val Phe Arg Gln Ala Gly Gly Ser Val Ala Leu Gly Ser Asp
305              310              315              320
Gln Ala Cys Gly Asn Asn Cys Cys Asn Ile Phe Asn Glu Met Lys Leu
      325              330              335
Thr Ala Leu Phe Asn Lys Ile Lys Tyr His Asp Pro Thr Ile Met Pro
      340              345              350
Ala Trp Glu Val Leu Arg Met Ala Thr Ile Glu Gly Ala Gln Ala Ile
      355              360              365
Gly Leu Asp His Lys Ile Gly Ser Leu Gln Val Gly Lys Glu Ala Asp
      370              375              380
Leu Ile Leu Ile Asp Leu Ser Ser Pro Asn Leu Ser Pro Thr Leu Leu
385              390              395              400
Asn Pro Ile Arg Asn Leu Val Pro Asn Leu Val Tyr Ala Ala Ser Gly
      405              410              415
His Glu Val Lys Ser Val Met Val Ala Gly Lys Leu Leu Val Glu Asp
      420              425              430
Tyr Gln Val Leu Thr Val Asp Glu Ser Ala Ile Leu Ala Glu Ala Gln
      435              440              445
Val Gln Ala Gln Gln Leu Cys Gln Arg Val Thr Ala Asp Pro Ile His
      450              455              460
Lys Lys Met Val Leu Met Glu Ala Met Ala Lys Gly Lys Leu
465              470              475

```

<210> 87

<211> 831

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 87

```

atgaccatat atgaaaataa gttgagtagt tatcaaaaaa atcaagatgc cataatatct      60
gcaaaagaac tcgaagaatg gcatttaatt ggacttctag accattcaat agatgcggta      120
atagtaccga attattttct tgagcaagag tgtatgacaa tttcagagag aataaaaaag      180
agtaaataatt ttagcgctta tcccgggtcat ccatcagtaa gtagcttggg acaagagttg      240
tatgaatgcg aaagtgcgct tgaattagca aagtatcaag aagacgcacc cacattgatt      300
aaagaaatgc ggaggctggt acatccgtac ataagtccaa ttgatagact taggggtgaa      360
gttgatgata tttggagtta tggctgtaat ttagcaaaac ttggtgataa aaaactgttt      420
gcgggtatcg ttagagagtt taaagaagat aaccctggcg caccacattg tgacgtaatg      480
gcatgggggtt ttctcgaata ttataaagat aaaccaaata tcataaatca aatcgagca      540
aatgtatatatt taaaaacgtc tgcacagga ggagaaatag tgctttggga tgaatggcca      600
actcaaagcg aatatatagc atacaaaaca gatgatccag ctagtcttcg tcttgatagc      660
aaaaagatcg cacaaccaa acttgagatc caaccgaacc agggagattt aattctattc      720
aattccatga gaattcatgc ggtgaaaaag atagaaactg gtgtacgtat gacatgggga      780
tgtttgattg gatactctgg aactgataaa ccgcttggtta tttggactta a      831

```

<210> 88

<211> 276

<212> PRT

15 <213> *Cylindrospermopsis raciborskii* AWT205

<400> 88

Met Thr Ile Tyr Glu Asn Lys Leu Ser Ser Tyr Gln Lys Asn Gln Asp
 1 5 10 15
 Ala Ile Ile Ser Ala Lys Glu Leu Glu Trp His Leu Ile Gly Leu
 20 25 30
 Leu Asp His Ser Ile Asp Ala Val Ile Val Pro Asn Tyr Phe Leu Glu
 35 40 45
 Gln Glu Cys Met Thr Ile Ser Glu Arg Ile Lys Lys Ser Lys Tyr Phe
 50 55 60
 Ser Ala Tyr Pro Gly His Pro Ser Val Ser Ser Leu Gly Gln Glu Leu
 65 70 75 80
 Tyr Glu Cys Glu Ser Glu Leu Glu Leu Ala Lys Tyr Gln Glu Asp Ala
 85 90 95
 Pro Thr Leu Ile Lys Glu Met Arg Arg Leu Val His Pro Tyr Ile Ser
 100 105 110
 Pro Ile Asp Arg Leu Arg Val Glu Val Asp Asp Ile Trp Ser Tyr Gly
 115 120 125
 Cys Asn Leu Ala Lys Leu Gly Asp Lys Lys Leu Phe Ala Gly Ile Val
 130 135 140
 Arg Glu Phe Lys Glu Asp Asn Pro Gly Ala Pro His Cys Asp Val Met
 145 150 155 160
 Ala Trp Gly Phe Leu Glu Tyr Tyr Lys Asp Lys Pro Asn Ile Ile Asn
 165 170 175
 Gln Ile Ala Ala Asn Val Tyr Leu Lys Thr Ser Ala Ser Gly Gly Glu
 180 185 190
 Ile Val Leu Trp Asp Glu Trp Pro Thr Gln Ser Glu Tyr Ile Ala Tyr
 195 200 205
 Lys Thr Asp Asp Pro Ala Ser Phe Gly Leu Asp Ser Lys Lys Ile Ala
 210 215 220
 Gln Pro Lys Leu Glu Ile Gln Pro Asn Gln Gly Asp Leu Ile Leu Phe
 225 230 235 240
 Asn Ser Met Arg Ile His Ala Val Lys Lys Ile Glu Thr Gly Val Arg
 245 250 255
 Met Thr Trp Gly Cys Leu Ile Gly Tyr Ser Gly Thr Asp Lys Pro Leu
 260 265 270
 Val Ile Trp Thr
 275

<210> 89

<211> 1398

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 89

10 ttaatgtagc gtttccattt gagtcaaggc acgagaagct tctaaagctg gaatagatac 60
 actatcattc tcaactacac tctcaaatgt cctaggtaac tgtgccccaa acatcagcat 120
 tccaatggcg ttgaacaaaa agaaagccaa ccacaagata tggttactct caaatTTAAC 180
 agcagctaca tccgcaggta aaaatcctac accaaacgcg attaagttaa cattgcggag 240
 agtatgccct tgagccaaac ccaagaagta cccacatagt atgcaacata ctgaattgca 300
 tactaggaca agtaccacac agggaataaa aatatcaata ttctcaataa tttctgcgtg 360
 gttggttaac aacccaaaaa catcatcggg aaatagccaa cacgctccgc cgaaaaccag 420
 actcactagc agagccattc ccacagaaac ttttgccaga ggtgctaact gttctgtggc 480
 tcctttccct ttaaaatttc ctgccagagt ttctgtacag aatcccaatc cttcaacaat 540
 gtagatgctc aaagcccata tctgtaagag caaggcattt tgagcgtaga taattgtccc 600
 catttgctgc ccttcgtagt taaacgttaa gttggtaaac atacaaacta aattgctgac 660

aaagatgttt ccattgagag ttaaggtgga gcgtatagct tttatgtccc aaatttttcc 720
 agctaattct ttacctctt gccacgggat ttctttgcag acaaaaaaca atcccaccaa 780
 tagggtgaga tattgacttg cagcagaagc tactcctgcc cccatgctcg accagtctaa 840
 gtggataata aacaagtagt cgagtgcat attggcagca ttgccacaa ccgacaacaa 900
 cacaactaag ccattttttt cccgtcccg aaaccagcca agcaggacaa agttgagcaa 960
 aatggcaggc gtcceccaac tctgggtgtt aaaatacgct tgagctgaag acttcacctc 1020
 tgggcccaga tctagtatag aaaaccccaa ccccctaac gggtagtga acagtatgat 1080
 cgccaccccc agcaccagag caattaaacc attagcagt cccgccaaca gtacgcctc 1140
 tcggtcatct cgtccgactg cttgtgctgt taacgcagtg gtaccattc gtaaaaacga 1200
 taaaacaaag tagagaaagt taagcaggtt tccagcaagg gctactccag ctaggtagtg 1260
 gatttccgag agatgaccta agaactgat actgactaaa ttactcagt gtactataat 1320
 attcgatagg acgttggtaa aagctagtcg gaagtagcgg ggtataaagt catactggct 1380
 tggaaatgtc aggtcat 1398

<210> 90

<211> 465
 <212> PRT
 <213> *Cylindrospermopsis raciborskii* AWT205

5 <400> 90

```

Met Ser Leu Thr Phe Pro Ser Gln Tyr Asp Phe Ile Pro Arg Tyr Phe
1      5      10      15
Arg Leu Ala Phe Thr Asn Val Leu Ser Asn Ile Ile Val Pro Leu Ser
      20      25      30
Asn Leu Val Ser Ile Met Phe Leu Gly His Leu Ser Glu Ile His Tyr
      35      40      45
Leu Ala Gly Val Ala Leu Ala Gly Asn Leu Leu Asn Phe Leu Tyr Phe
      50      55      60
Val Leu Ser Phe Leu Arg Met Gly Thr Thr Ala Leu Thr Ala Gln Ala
65      70      75      80
Val Gly Arg Asp Asp Arg Glu Gly Val Leu Leu Ala Gly Leu Leu Asn
      85      90      95
Gly Leu Ile Ala Leu Val Leu Gly Val Ala Ile Ile Leu Leu Gln Tyr
      100     105     110
Pro Leu Gly Val Leu Gly Phe Ser Ile Leu Asp Val Gly Pro Glu Val
      115     120     125
Lys Ser Ser Ala Gln Ala Tyr Phe Asn Thr Gln Ser Trp Gly Ala Pro
      130     135     140
Ala Ile Leu Leu Asn Phe Val Leu Leu Gly Trp Phe Leu Gly Arg Glu
145     150     155     160
Lys Asn Gly Leu Val Val Leu Leu Ser Val Val Gly Asn Ala Ala Asn
      165     170     175
Ile Ala Leu Asp Tyr Leu Phe Ile Ile His Leu Asp Trp Ser Ser Met
      180     185     190
Gly Ala Gly Val Ala Ser Ala Ala Ser Gln Tyr Leu Thr Leu Leu Val
      195     200     205
Gly Leu Phe Phe Val Cys Lys Glu Ile Pro Trp Gln Glu Val Lys Glu
      210     215     220
Leu Ala Gly Lys Ile Trp Asp Ile Lys Ala Ile Arg Ser Thr Leu Thr
225     230     235     240
Leu Asn Gly Asn Ile Phe Val Ser Asn Leu Val Cys Met Phe Thr Asn
      245     250     255
Leu Thr Phe Asn Tyr Glu Gly Ala Gln Met Gly Thr Ile Ile Tyr Ala
      260     265     270
Gln Asn Ala Leu Leu Leu Gln Ile Trp Ala Leu Ser Ile Tyr Ile Val
      275     280     285
Glu Gly Leu Gly Phe Cys Thr Glu Thr Leu Ala Gly Asn Phe Lys Gly
290     295     300

Lys Gly Ala Thr Glu Gln Leu Ala Pro Leu Ala Lys Val Ser Val Gly
305     310     315     320
Met Ala Leu Leu Val Ser Leu Val Phe Gly Gly Ala Cys Trp Leu Phe
      325     330     335
Pro Asp Asp Val Phe Gly Leu Leu Thr Asn His Ala Glu Ile Ile Glu
      340     345     350
Asn Ile Asp Ile Phe Ile Pro Trp Leu Val Leu Val Leu Val Cys Asn
      355     360     365
Ser Val Cys Cys Ile Leu Cys Gly Tyr Phe Leu Gly Leu Ala Gln Gly
      370     375     380
His Thr Leu Arg Asn Val Asn Leu Ile Ala Phe Gly Val Gly Phe Leu
385     390     395     400
Pro Ala Asp Val Ala Ala Val Lys Phe Glu Ser Asn His Ile Leu Trp
      405     410     415
Leu Ala Phe Phe Leu Phe Asn Ala Ile Gly Met Leu Met Phe Gly Ala
      420     425     430
Gln Leu Pro Arg Thr Phe Glu Ser Val Val Glu Asn Asp Ser Val Ser
      435     440     445
Ile Pro Ala Leu Glu Ala Ser Arg Ala Leu Thr Gln Met Glu Thr Leu
450     455     460
His
465

```

10

<210> 91
 <211> 750

<212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 91

5

```

atgttgaact tagaccgcat cctgaatcaa gagcgactgc tacgagaaat gactggactt      60
aaccgccaag cattcaacga gctgttatct cagtttgctg atacctatga acgcaccgtg      120
ttcaactcct tagcaaaccg caaacgtgcg cccgggggcg gacgcaagcc tacactcaga      180
agtatagagg aaaaactatt ttatatcctg ctgtactgca aatgttatcc gacgtttgac      240
ttgctgagtg tgttggttcaa ctttgaccgc tcctgtgctc atgattgggt acatcgacta      300
ctgtctgtgc tagaaaccac tttaggagaa aagcaagttt tgccagcacg caaactcagg      360
agcatggagg aattcaccaa aaggtttcca gatgtgaagg aggtgattgt ggatggtacg      420
gagcgtccag tccagcgtcc tcaaaaccga gaacgccaaa aagagtatta ctctggcaag      480
aaaaagcggc atacatgcaa gcagattaca gtcagcaciaa gggagaaacg agtgattatt      540
cggacggaaa ccagagcagg taaagtgcac gacaaacggc tactccatga atcagagata      600
gtgcaataca ttcctgatga agtagcaata gagggagatt tgggttttca tgggttggag      660
aaagaatttg tcaatgtcca tttaccacac aagaaccga aaggtatcga agcaaggagg      720
catggcggcg ggatgggtca gtttttataa      750

```

<210> 92

<211> 249

10

<212> PRT

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 92

15

```

Met Leu Asn Leu Asp Arg Ile Leu Asn Gln Glu Arg Leu Leu Arg Glu
1      5      10      15
Met Thr Gly Leu Asn Arg Gln Ala Phe Asn Glu Leu Leu Ser Gln Phe
20     25     30
Ala Asp Thr Tyr Glu Arg Thr Val Phe Asn Ser Leu Ala Asn Arg Lys
35     40     45
Arg Ala Pro Gly Gly Gly Arg Lys Pro Thr Leu Arg Ser Ile Glu Glu
50     55     60
Lys Leu Phe Tyr Ile Leu Leu Tyr Cys Lys Cys Tyr Pro Thr Phe Asp
65     70     75     80
Leu Leu Ser Val Leu Phe Asn Phe Asp Arg Ser Cys Ala His Asp Trp
85     90     95
Val His Arg Leu Leu Ser Val Leu Glu Thr Thr Leu Gly Glu Lys Gln
100    105    110
Val Leu Pro Ala Arg Lys Leu Arg Ser Met Glu Glu Phe Thr Lys Arg
115    120    125
Phe Pro Asp Val Lys Glu Val Ile Val Asp Gly Thr Glu Arg Pro Val
130    135    140
Gln Arg Pro Gln Asn Arg Glu Arg Gln Lys Glu Tyr Tyr Ser Gly Lys
145    150    155    160
Lys Lys Arg His Thr Cys Lys Gln Ile Thr Val Ser Thr Arg Glu Lys
165    170    175
Arg Val Ile Ile Arg Thr Glu Thr Arg Ala Gly Lys Val His Asp Lys
180    185    190
Arg Leu Leu His Glu Ser Glu Ile Val Gln Tyr Ile Pro Asp Glu Val
195    200    205
Ala Ile Glu Gly Asp Leu Gly Phe His Gly Leu Glu Lys Glu Phe Val
210    215    220
Asn Val His Leu Pro His Lys Lys Pro Lys Gly Ile Glu Ala Arg Arg
225    230    235    240
His Gly Gly Gly Met Gly Gln Phe Leu
245

```

<210> 93

20

<211> 1431

<212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 93

25

atgaatctta	taacaacaaa	aaaacaggta	gatacattag	tgatacacgc	tcattcttttt	60
accatgcagg	gaaatggtgt	gggatatatt	gcagatgggg	cacttgcggt	tgagggttagc	120
cgtattgtag	cagttgattc	gacggaggcg	ttgctgagtc	atcttgaggg	cagaaagggtt	180
attgagtcgg	cgaattgtgc	cgtcttgcc	gggctgatta	atgctcacgt	agacacaagt	240
ttggtgctga	tgcgtggggc	ggcgcaagat	gtaactaatt	ggctaattga	cgcgaccatg	300
ccttattttg	ctcacatgac	acccgtggcg	agtatggctg	caacacgctt	aagggtggta	360
gaagagttag	aagcaggcac	aacaacattc	tgtgacaata	aaattattag	ccccctgtgg	420
ggcgaaat	tcgatgaaat	tggtgtacgg	gctagtttag	ctcctatggt	cgatgcactc	480
ccactggaga	tgccaccgct	tcaagacggg	gagctttatc	ccttcgatat	caaggcgagg	540
cggcgggcga	tggcagaggc	tgtggatttt	gcctgtgggt	ggaatggggc	agcagagggg	600
cgtatcacta	ccatgttagg	aatgtattcg	ccagatatga	tgccgcttga	gatgctacgc	660
gcagccaaag	agattgctca	acgggaaggc	ttaatgctgc	atcttcattgt	agcgcaggga	720
gatcgggaaa	cagagcaaat	cgttaaacga	tatggtaagc	gtccgatcgc	atctttagct	780
gagattggct	acttggacga	acagttgctg	gcagttcacc	tcaccgatgc	caccgatgaa	840
gaggtgatac	aagtagccaa	aagtggcgct	ggcatggtag	tctgttcggg	aatgattggc	900
actattgacg	gtatcgtgcc	gcccgtcat	gtgtttcggc	aagcaggcgg	acccgttgcg	960
ctaggcagca	gctacaataa	tattttccat	gagatgaagc	tgaccgcctt	attcaacaaa	1020
ataaaatc	acgatccaac	cattatgccg	gcttgggaag	tcctgcgtat	ggctaccatc	1080
gaaggagcgc	ggcgattggg	tttagatcac	aagattggct	ctcttgaagt	tggcaaagaa	1140
gccgacctga	tcttaataga	cctcagcacc	cctaacctct	cacccactct	gcttaacccc	1200
attcgttaacc	ttgtaccta	tttcgtgtac	gctgcttcag	gacatgaagt	taaaagtgtc	1260
atggtggcgg	gaaaactgtt	attggaagac	taccaagtcc	tcacagtaga	tgagtctgct	1320
atcattgctg	aagcacaatt	gcaagcccaa	cagatttctc	aatgcgtagc	atctgaccct	1380
atccacaaaa	aaatggtgct	gatggcgggc	atggcaaggg	gccaattgta	g	1431

<210> 94

<211> 476

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 94

```

Met Asn Leu Ile Thr Thr Lys Lys Gln Val Asp Thr Leu Val Ile His
1      5      10      15
Ala His Leu Phe Thr Met Gln Gly Asn Gly Val Gly Tyr Ile Ala Asp
      20      25      30
Gly Ala Leu Ala Val Glu Gly Ser Arg Ile Val Ala Val Asp Ser Thr
      35      40      45
Glu Ala Leu Leu Ser His Phe Glu Gly Arg Lys Val Ile Glu Ser Ala
      50      55      60
Asn Cys Ala Val Leu Pro Gly Leu Ile Asn Ala His Val Asp Thr Ser
65      70      75      80
Leu Val Leu Met Arg Gly Ala Ala Gln Asp Val Thr Asn Trp Leu Met
      85      90      95
Asp Ala Thr Met Pro Tyr Phe Ala His Met Thr Pro Val Ala Ser Met
      100     105     110
Ala Ala Thr Arg Leu Arg Val Val Glu Glu Leu Lys Ala Gly Thr Thr
      115     120     125
Thr Phe Cys Asp Asn Lys Ile Ile Ser Pro Leu Trp Gly Glu Phe Phe
      130     135     140
Asp Glu Ile Gly Val Arg Ala Ser Leu Ala Pro Met Phe Asp Ala Leu
145     150     155     160
Pro Leu Glu Met Pro Pro Leu Gln Asp Gly Glu Leu Tyr Pro Phe Asp
      165     170     175
Ile Lys Ala Gly Arg Arg Ala Met Ala Glu Ala Val Asp Phe Ala Cys
      180     185     190
Gly Trp Asn Gly Ala Ala Glu Gly Arg Ile Thr Thr Met Leu Gly Met
      195     200     205
Tyr Ser Pro Asp Met Met Pro Leu Glu Met Leu Arg Ala Ala Lys Glu
      210     215     220
Ile Ala Gln Arg Glu Gly Leu Met Leu His Phe His Val Ala Gln Gly
225     230     235     240
Asp Arg Glu Thr Glu Gln Ile Val Lys Arg Tyr Gly Lys Arg Pro Ile
      245     250     255
Ala Phe Leu Ala Glu Ile Gly Tyr Leu Asp Glu Gln Leu Leu Ala Val
      260     265     270
His Leu Thr Asp Ala Thr Asp Glu Val Ile Gln Val Ala Lys Ser
      275     280     285
Gly Ala Gly Met Val Leu Cys Ser Gly Met Ile Gly Thr Ile Asp Gly
      290     295     300
Ile Val Pro Pro Ala His Val Phe Arg Gln Ala Gly Gly Pro Val Ala
305     310     315     320
Leu Gly Ser Ser Tyr Asn Asn Ile Phe His Glu Met Lys Leu Thr Ala
      325     330     335
Leu Phe Asn Lys Ile Lys Tyr His Asp Pro Thr Ile Met Pro Ala Trp
      340     345     350
Glu Val Leu Arg Met Ala Thr Ile Glu Gly Ala Arg Ala Ile Gly Leu
      355     360     365
Asp His Lys Ile Gly Ser Leu Glu Val Gly Lys Glu Ala Asp Leu Ile
      370     375     380
Leu Ile Asp Leu Ser Thr Pro Asn Leu Ser Pro Thr Leu Leu Asn Pro
385     390     395     400
Ile Arg Asn Leu Val Pro Asn Phe Val Tyr Ala Ala Ser Gly His Glu
      405     410     415
Val Lys Ser Val Met Val Ala Gly Lys Leu Leu Leu Glu Asp Tyr Gln
      420     425     430
Val Leu Thr Val Asp Glu Ser Ala Ile Ile Ala Glu Ala Gln Leu Gln
      435     440     445

Ala Gln Gln Ile Ser Gln Cys Val Ala Ser Asp Pro Ile His Lys Lys
450     455     460
Met Val Leu Met Ala Ala Met Ala Arg Gly Gln Leu
465     470     475

```

5 <210> 95
 <211> 780
 <212> ADN
 <213> Cylindrospermopsis raciborskii AWT205

10 <400> 95

```

atgcaagaaa aacgaatcgc aatgtggtct gtgccacgaa gtttgggtac agtgctgcta 60
caagcctggg cgagtcggcc agataccgta gtctttgatg aacttctctc ctttccctat 120
ctctttatca aaggggaaaga tatgggcttt acttggacag accttgattc tagccaaatg 180
ccccacgcag attggcgatc cgtcatcgat ctgttaaagg ctcccctgcc tgaagggaaa 240
tcaatcatcg atctgttaaa ggctcccctg cctgaaggga aatcaatttg ctatcagaag 300
catcaagcgt atcatttaat cgaagagacc atggggattg agtggatatt gcccttcagc 360
aactgctttc tgattcgcca acccaaagaa atgctcttat cttttcgtaa gattgtgcca 420
cattttacct ttgaagaaac aggcctggatc gaattaaaac ggctgtttga ctatgtacat 480
caaacgagcg gagtaatccc gcctgtcata gatgcacacg acttgctgaa cgatccgcgg 540
agaatgctct ccaagctttg tcagggttga ggggttgagt ttaccgagac aatgctcagt 600
tggtccccc tggaggtcga gttgaacgaa aaactagccc cttggtacag caccgtagca 660
agttctacgc attttcactc gtatcagaat aaaaatgagt cgttgccgct atatcttgtc 720
gatatttgta aacgctgcga tgaaatatat caggaattat atcaatttcg actttattag 780

```

<210> 96
 <211> 259
 5 <212> PRT
 <213> *Cylindrospermopsis raciborskii* AWT205

<400> 96

```

Met Gln Glu Lys Arg Ile Ala Met Trp Ser Val Pro Arg Ser Leu Gly
1      5      10      15
Thr Val Leu Leu Gln Ala Trp Ser Ser Arg Pro Asp Thr Val Val Phe
      20      25      30
Asp Glu Leu Leu Ser Phe Pro Tyr Leu Phe Ile Lys Gly Lys Asp Met
      35      40      45
Gly Phe Thr Trp Thr Asp Leu Asp Ser Ser Gln Met Pro His Ala Asp
      50      55      60
Trp Arg Ser Val Ile Asp Leu Leu Lys Ala Pro Leu Pro Glu Gly Lys
65      70      75      80
Ser Ile Ile Asp Leu Leu Lys Ala Pro Leu Pro Glu Gly Lys Ser Ile
      85      90      95
Cys Tyr Gln Lys His Gln Ala Tyr His Leu Ile Glu Glu Thr Met Gly
      100     105     110
Ile Glu Trp Ile Leu Pro Phe Ser Asn Cys Phe Leu Ile Arg Gln Pro
      115     120     125
Lys Glu Met Leu Leu Ser Phe Arg Lys Ile Val Pro His Phe Thr Phe
      130     135     140
Glu Glu Thr Gly Trp Ile Glu Leu Lys Arg Leu Phe Asp Tyr Val His
145     150     155     160
Gln Thr Ser Gly Val Ile Pro Pro Val Ile Asp Ala His Asp Leu Leu
      165     170     175
Asn Asp Pro Arg Arg Met Leu Ser Lys Leu Cys Gln Val Val Gly Val
      180     185     190
Glu Phe Thr Glu Thr Met Leu Ser Trp Pro Pro Met Glu Val Glu Leu
      195     200     205
Asn Glu Lys Leu Ala Pro Trp Tyr Ser Thr Val Ala Ser Ser Thr His
10      210     215     220

Phe His Ser Tyr Gln Asn Lys Asn Glu Ser Leu Pro Leu Tyr Leu Val
225     230     235     240
Asp Ile Cys Lys Arg Cys Asp Glu Ile Tyr Gln Glu Leu Tyr Gln Phe
      245     250     255
Arg Leu Tyr

```

<210> 97
 15 <211> 1176
 <212> ADN
 <213> *Cylindrospermopsis raciborskii* AWT205

<400> 97

```

atgcaaacaa gaattgtaaa tagctggaat gagtgggatg aactaaagga gatgggtgtc 60
gggattgcag atgggtgctta ttttgaacca actgagccag gtaaccgccc tgctttacgc 120
gataagaaca ttgccaaaat gttctctttt ccagggggtc cgaaaaagca agaggtaaca 180
gagaaagcta atgaggagtt gaatgggctg gtagcgcttc tagaatcaca gggcgtaact 240
gtacgccgcc cagagaaaaca taactttggc ctgtctgtga agacaccatt ctttgaggta 300
gagaatcaat attgtgcggt ctgccacgt gatgttatga tcacctttgg gaacgaaatt 360
ctcgaagcaa ctatgtcacg gcggtcacgc ttctttgagt atttacccta tcgcaaacta 420
gtctatgaat attggcataa agatccagat atgatctgga atgctgcgcc taaaccgact 480
atgcaaaatg ccatgtaccg cgaagatttc tgggagtgtc cgatggaaga tcgatttgag 540
agtatgcatg attttgagtt ctgctgcacc caggatgagg tgatttttga cgcagcagac 600
tgtagccgct ttggccgtga tatttttgtg caggagtcaa tgacgactaa tcgtgcaggg 660
attcgtggc tcaaaccgca tttagagccg cgctgcttcc gcgtgcatga tattcacttc 720
ccactagata ttttcccatc ccacattgat tgtacttttg tccccttagc acctgggggtt 780
gtgttagtga atccagatcg ccccatcaaa gaggggtgaag agaaactctt catggataac 840
ggttggcaat tcatacgaagc accctcccc acttccaccg acgatgagat gcctatgttc 900
tgccagtcca gtaagtgggt ggcatgaat gtgttaagca tttcccccga gaaggatc 960
tgtgaagagc aagagcatcc gttcatgag ttgctagata aacacggctt tgagggtctat 1020
ccaattccct ttgcgaatgt ctttgagttt ggcgggttcgc tccattgtgc cacctgggat 1080
atccatcgca cgggaacctg tgaggattac ttccctaaac taaactatac gccggttaact 1140
gcataacca atggcggttc tcgcttcac attttag 1176

```

<210> 98

<211> 391

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 98

```

Met Gln Thr Arg Ile Val Asn Ser Trp Asn Glu Trp Asp Glu Leu Lys
1      5      10      15
Glu Met Val Val Gly Ile Ala Asp Gly Ala Tyr Phe Glu Pro Thr Glu
20     25     30
Pro Gly Asn Arg Pro Ala Leu Arg Asp Lys Asn Ile Ala Lys Met Phe
35     40     45
Ser Phe Pro Arg Gly Pro Lys Lys Gln Glu Val Thr Glu Lys Ala Asn
50     55     60
Glu Glu Leu Asn Gly Leu Val Ala Leu Leu Glu Ser Gln Gly Val Thr
65     70     75     80
Val Arg Arg Pro Glu Lys His Asn Phe Gly Leu Ser Val Lys Thr Pro
85     90     95
Phe Phe Glu Val Glu Asn Gln Tyr Cys Ala Val Cys Pro Arg Asp Val
100    105    110
Met Ile Thr Phe Gly Asn Glu Ile Leu Glu Ala Thr Met Ser Arg Arg
115    120    125
Ser Arg Phe Phe Glu Tyr Leu Pro Tyr Arg Lys Leu Val Tyr Glu Tyr
130    135    140
Trp His Lys Asp Pro Asp Met Ile Trp Asn Ala Ala Pro Lys Pro Thr
145    150    155    160

```

10

Met Gln Asn Ala Met Tyr Arg Glu Asp Phe Trp Glu Cys Pro Met Glu
165 170 175
Asp Arg Phe Glu Ser Met His Asp Phe Glu Phe Cys Val Thr Gln Asp
180 185 190
Glu Val Ile Phe Asp Ala Ala Asp Cys Ser Arg Phe Gly Arg Asp Ile
195 200 205
Phe Val Gln Glu Ser Met Thr Asn Arg Ala Gly Ile Arg Trp Leu
210 215 220
Lys Arg His Leu Glu Pro Arg Arg Phe Arg Val His Asp Ile His Phe
225 230 235 240
Pro Leu Asp Ile Phe Pro Ser His Ile Asp Cys Thr Phe Val Pro Leu
245 250 255
Ala Pro Gly Val Val Leu Val Asn Pro Asp Arg Pro Ile Lys Glu Gly
260 265 270
Glu Glu Lys Leu Phe Met Asp Asn Gly Trp Gln Phe Ile Glu Ala Pro
275 280 285
Leu Pro Thr Ser Thr Asp Asp Glu Met Pro Met Phe Cys Gln Ser Ser
290 295 300
Lys Trp Leu Ala Met Asn Val Leu Ser Ile Ser Pro Lys Lys Val Ile
305 310 315 320
Cys Glu Glu Gln Glu His Pro Leu His Glu Leu Leu Asp Lys His Gly
325 330 335
Phe Glu Val Tyr Pro Ile Pro Phe Arg Asn Val Phe Glu Phe Gly Gly
340 345 350
Ser Leu His Cys Ala Thr Trp Asp Ile His Arg Thr Gly Thr Cys Glu
355 360 365
Asp Tyr Phe Pro Lys Leu Asn Tyr Thr Pro Val Thr Ala Ser Thr Asn
370 375 380
Gly Val Ser Arg Phe Ile Ile
385 390

<210> 99

<211> 8754

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 99

atgcaaaaga gagaaagccc acagatacta tttgatggga atggaacaca atctgagttt 60
ccagatagtt gcattcacca cttgttcgag gatcaagccg caaagcgacc ggatgcgac 120
gctctcattg acggtgagca atcccttacc tacggggaac taaatgtacg cgctaaccac 180
ctagcccagc atctcttgte cctaggctgt caacccgatg acctcctcgc catctgcatc 240
gagcgttcgg cagaactcct tattgggtttg ttgggtatcc taaaagccgg atgtgcttat 300
gtgcctttgg atgtaggcta tcctggcgat cgcataagat atatgttgcg ggactcggat 360
gcgcgtatct tactaacctc aacggatgtc gctaagaaac ttgccttaac catacctgca 420
ttgcaagagt gccaaaccgt ctatttagat caagagatat ttgagtatga ttttcatttt 480
ttagcgatag cttaaactatt acataaccaa tacttgagat tattacattt ttatttttat 540
accttgatcc agcaatgccg ggcaacttcg gtttcccaag ggattcagac acagggttctc 600
cccaataatc tcgcttactg catctacacc tctggctcta ccggaatcc caaagggatc 660
ttgatgggaa atcgctcact ggtgaatatg ctttgggtgg atcagcaaac gcggccttcg 720
gttcagggtg ttaggacgct gcaattttgt gcagtcagct ttgacttttc ctgccatgaa 780
atcttttcta ccctctgtct tggcgggata ttggcttggg tgccagaggc agtgcgcaa 840
aatccctttg cattggctga gttcatcagt caacagaaaa ttgaaaaatt gtttcttccc 900
gttatagcat tactacagtt ggccgaagct gtaaattggga ataaaagcac ctccctcgcg 960
ctttgcgaag ttatcactac cggggagcag atgcagatca cacctgctgt cgccaacctc 1020
tttcagaaaa ccggggcgat gttgcataat cactacgggg caacagaatt tcaagatgcc 1080
accactcata ccctcaaggg caatccagag ggctggccaa cactggtgcc agtgggtcgt 1140
ccactgcaca atgttcaagt gtatattctg gatgaggcac agcaacctgt acctcttgg 1200
ggagagggtg aattctgtat tgggtgtatt ggactggctc gtggctatca caatttgcct 1260

10

gacctaacga	atgaaaaatt	tattcccaat	ccatttgggg	ctaattgagaa	cgctaaaaaa	1320
ctctaccgca	caggggactt	ggcacgctac	ctacccgacg	gcacgattga	gcatttagga	1380
cggatagacc	accagggttaa	gatccgaggt	ttccgcgtgg	aattggggga	aattgagtc	1440
gtgctggcaa	gtcaccaagc	tgtgcgtgaa	tgtgccgttg	tggcacggga	gattgcaggt	1500
catacacagt	tggtagggtta	tatcatagca	aaggatacac	ttaattctcag	tttcgacaaa	1560
cttgaacctta	tcctgcgtca	atattcggaa	gcggtgctgc	cagaatacat	gataccact	1620
cggttcatca	atatcagtaa	tatgccgttg	actcccagtg	gtaaaacttga	ccgcagggca	1680
ttacctgatc	ccaaaggcga	tcgccctgca	ttgtctaccc	cacttgtcaa	gcctcgtacc	1740
cagacagaga	aacgtttagc	agagatttgg	ggcagttatc	ttgctgtaga	tattgtggga	1800
acccacgaca	atttctttga	tctaggcggg	acgtcactgc	tattgactca	agcgacacaa	1860
ttcctgtgcg	agacctttaa	tattaatttg	tccgctgtct	cactctttca	atatcccaca	1920
attcagacat	tgccacaata	tattgattgc	caaggagaca	caacctcaag	cgatacagca	1980
tccaggcaca	agaaagtacg	taaaaagcag	tccgggtgaca	gcaacgatat	tgccatcatc	2040
agtgtggcag	gtcgttttcc	gggtgctgaa	acgattgagc	agttctggca	taatctctgt	2100
aatggtggtg	aatccatcac	cctttttagt	gatgatgagc	tagagcagac	tttgccctgag	2160
ttatttaata	atcccgcctta	tgtcaaagca	gggtgcgggtgc	tagaaggcgt	tgaattattt	2220
gatgtcacct	tttttggtca	cagcccaaaa	gaagctgcgg	tgacagaccc	tcagcaacgg	2280
attttgctag	agtgtgctcg	ggaagcattt	gaacgggctg	gctacaaccc	cgaaacctat	2340
ccagaaccag	ttgggtgtta	tgctggttca	agcctgagta	cctatctgct	taacaatatt	2400
ggctctgctt	taggcataat	taccgagcaa	ccctttattg	aaacggatat	ggagcagttt	2460
caggctaaaa	ttggcaatga	ccggagctat	cttgctacac	gcactcttta	caagctgaat	2520
ctcaagggtc	caagcgctca	tgtgcagacc	gcctgctcaa	cctcgttagt	tgccggtcac	2580
atggcctgtc	agagtctcat	tagtgagag	tgtcaaatgg	ctttagccgg	tggtatttct	2640
gtggttgtag	cacagaaggg	gggctatctc	tacgaagaag	gcatggttcg	ttccaggat	2700
ggctcatgtc	gcgcctttga	tgccgaagcc	caagggacta	tatttggcaa	tggcggcgcc	2760
ttggttttgc	ttaaacggtt	gcaggatgca	ctggacgata	acgacaacat	tatggcagtc	2820
atcaaagcca	cagccatcaa	caacgacggt	gcgctcaaga	tgggctacac	agcaccgagc	2880
gtggatgggc	aagctgatgt	aattagcgag	gcgattgcta	tcgctgacat	agatgcaagc	2940
accattggct	atgtagaagc	tcattggcaca	gccacccaat	tgggtgatcc	gattgaahta	3000
gcagggttag	caagggcatt	tcagcgtagt	acggacagcg	tccttggtta	acaacaatgc	3060
gctattggat	cagttaaaac	taataattggc	cacttagatg	aggcggcagg	cattgccgga	3120
ctgataaagg	ctgctctagc	tctacaatat	ggacagattc	caccgagctt	gcactatgcc	3180
aatcctaata	cacggattga	ttttgacgca	acccattttt	ttgtcaacac	agaactacgc	3240
gaatgggtcaa	ggaatggtta	tcctcggcgg	gcgggggtga	gttcttttgg	tgtgggtgga	3300
actaacagcc	atattgtgct	ggaggagtcg	cctgtaaaagc	aaccacattt	gttctcttct	3360
ttgccagaac	gcagtcatca	tctgctgacg	ctttctgccc	atacacaaga	ggctttgcat	3420
gagttggtgc	aacgctacat	ccaacataac	gagacacacc	ttgatattaa	cttaggcgac	3480
ctctgtttca	cagccaatac	gggacgcaag	cattttgagc	atcgctagc	ggttgtagcc	3540
gaatcaaatcc	ctggccttaca	ggcacaactg	gaaactgcac	agactgcgat	ttcagcacag	3600
aaaaaaaaatg	ccccgccgac	gatcgcatte	ctgtttacag	gtcaaggctc	acaatacatt	3660
aacatggggc	gcaccctcta	cgatactgaa	tcaacattcc	gtgcagccct	tgaccgatgt	3720
gaaaccattc	tccaaaattt	agggatcgag	tccattctct	ccgttatttt	tggttcatct	3780
gagcatggac	tctcattaga	tgacacagcc	tatacccagc	ccgcactctt	tgccatcgaa	3840
tacgcgctct	atcaattatg	gaagtcgtgg	ggcatccagc	cctcagtggt	gataggtcat	3900
agtgtagggt	aatatgtgtc	cgcttgtgtg	gcgggagtc	ttagcttaga	ggatgggttg	3960
aaactgattg	cagaacgagg	acgactgata	caggcacttc	ctcgtgatgg	gagcatggtt	4020
tccgtgatgg	caagcgagaa	gcgtattgca	gatatcattt	taccttatgg	gggacaggta	4080
gggatcgccg	cgattaatgg	cccacaaagt	gttgtaattt	ctgggcaaca	gcaagcgatt	4140
gatgctatct	gtgccatctt	ggaaactgag	ggcatcaaaa	gcaagaagct	aaacgtctcc	4200
catgccttcc	actcgccgct	agtggaagca	atggttagact	ctttcttgca	ggttgcacaa	4260
gaggtcactt	actcgcaacc	tcaaatcaag	cttatctcta	atgtaacggg	aacattggca	4320
agccatgaat	cttgctccga	tgaacttccg	atcaccaccg	cagagtattg	ggtacgtcat	4380
gtgcgacagc	ccgtccggtt	tgcggcgga	atggagagcc	ttgaggggtca	aggggtaaac	4440
gtatttatag	aaatcgggtc	taaacctgtt	cttttaggca	tgggacgcga	ctgcttgcc	4500
gaacaagagg	gacttttggt	gcctagtttg	cgcccaaac	aggatgattg	gcaacaggtg	4560
ttaagtagtt	tgcgtgatct	atacttagca	ggtgtaaccg	tagattggag	cagtttcgat	4620
caggggtatg	ctcgtcgccg	tgtgccacta	ccgacttatc	cttggcagcg	agagcgccat	4680

tgggtagagc	caattattcg	tcaacggcaa	tcagtattac	aagccacaaa	taccaccaag	4740
ctaactcgta	acgccagcgt	ggcgagcat	cctctgcttg	gtcaacggct	gcatttgtcg	4800
cggactcaag	agatttactt	tcaaaccttc	atccactccg	acttcccaat	atgggttgct	4860
gatcataaag	tatttggaag	tgtcatcatt	ccgggtgtcg	cctattttga	gatggcactg	4920
gcagcagggg	aggcacttaa	accagacagt	atattttggc	tcgaagatgt	atccatcgcc	4980
caagcactga	ttattcccga	tgaaggggcaa	actgtgcaaa	tagtattaag	cccacaggaa	5040
gagtcagctt	atTTTTTTga	aatcctctct	ttagaaaaag	aaaactcttg	ggtgcttcat	5100
gcctctggta	agctagtcgc	ccaagagcaa	gtgctagaaa	ccgagccaat	tgacttgatt	5160
gcgttacagg	cacattgttc	cgaagaagtg	tcagtagatg	tgctatatca	ggaagaaatg	5220
gcgcgcgggc	tggatatggg	tccaatgatg	cgtgggggtg	agcagctttg	gcgttatccg	5280
ctctcctttg	ccaaaagtca	tgatcgcatc	gcactcgcca	aggtcagctt	gccagaaatc	5340
ttgcttcatg	agtccaatgc	ctaccaattc	catcctgtaa	tcttggtatg	ggggctgcaa	5400
atgataacgg	tctcttatcc	tgaagcaaac	caaggccaga	cttatgtacc	tggttggtata	5460
gagggtctac	aagtctatgg	tcgtcccagt	tcagaacttt	ggtgtcgcg	ccaatatcgg	5520
cctccttttg	atacagatca	aaggcagggt	attgatttgc	tgccaaagaa	attgattgca	5580
gacttgcatc	tatttgatac	ccagggtcgt	gtggttgcca	tcattgtttg	tggtcaatct	5640
gtccttgttg	gacgggaagc	aatgttgcca	tcgcaagata	cttgccgaaa	ttggctttat	5700
caagtccctg	ggaaacctca	agcctgtttt	ggacttttac	cgaattacct	gccaaaccca	5760
gataagattc	ggaaacgcct	ggaaacaaag	ttagcgacat	tgatcatcga	agctaatttg	5820
gcgacttatg	cgatcgccca	tacccaactg	gaaagggtta	gtctagctta	cgttggtggc	5880
gctttccgac	aaatgggctg	gctgtttcaa	cccggtgagc	gtttttccac	cgccagaaag	5940
gtatcagcgt	taggaatcgt	tgatcaacat	cggcaactat	tcgctcgttt	gctcgacatt	6000
ctagccgaag	cagacatact	ccgcagcgaa	aacttgatga	cgatatggga	agtcatttca	6060
tacccggaag	cgattgatat	acaggtaact	cttgacgacc	tcgaagccaa	agaagcagaa	6120
gccgaagtca	cactggtttc	ccgttgacgt	gcaaaattgg	ccgaagtatt	acaaggaaaa	6180
tgtgacccca	tacagttgct	ctttcccgca	ggggacacaa	caacgttaag	caactcttat	6240
cgtgaagccc	cagttttggg	tgttactaat	actctagtcc	aagaagcgct	tctttccgcc	6300
ctggagcagt	tgccgcggga	acgtgggttg	cgaattttag	agattgggtg	tggaacaggt	6360
ggaaccacag	cctacttggt	accgcatctg	cctggggatc	agacaaaata	tgcttttacc	6420
gatattagtg	ccttttttct	tgccaaagcg	gaagagcggt	ttaaagatta	cccgtttgta	6480
gcgttatcag	tattagatat	cgaacaagca	ccagggcg	aaggatttga	acccaaata	6540
tacgatttaa	tcgtagcagc	ggatgtcttg	catgctacta	gtgacctg	tcaaactctt	6600
gtacatatcc	ggcaattatt	agcgccgggc	gggatgttga	tcctgatgga	agacagcgaa	6660
ccgcacgct	gggctgattt	aacctttggc	ttaacagaag	gctggtggaa	gtttacagac	6720
catgacttac	gccccaaaca	tccgctattg	tctctgagc	agtggcaaat	cttggtgtca	6780
gaaatgggag	ttagtcaaac	aaccgcctta	tgcccaaaaa	tagatagccc	ccataaattg	6840
ccacgggagg	cgggtgattg	ggcgcgtaat	gaaccagcca	tcagaaaacc	ccgaagatgg	6900
ctgatcttgg	ctgacgagga	gattgggtgga	ctactagcca	aacagctacg	tgaagaagga	6960
gaagattgta	tactcctctt	gccaggggaa	aagtacacag	agagagattc	acaaacgttt	7020
acaatcaatc	ctggagatat	tgaagagtgg	caacagttat	tgaaccgagt	accgaacata	7080
caagaaattg	tacattgttg	gagtatggtt	tccactgact	tagatagagc	cactattttc	7140
agttgcagca	gtacgctgca	tttagttcaa	gcattagcaa	actatccaaa	aaacctcgc	7200
ttgtcacttg	tcaccctagg	cgcacaagcc	gttaacgaac	atcatgttca	aaatgtagtt	7260
ggagcagccc	tctggggcat	gggaaaggta	attgcaactg	aacaccagga	gctacaagta	7320
gcacaaatgg	atttagaccc	gaatgggaag	gttaaggcgc	aagtagaagt	gcttagggat	7380
gaactttctg	ccagaaaaaga	ccctgcatca	gcaatgtctg	tgccctgatc	gcaaacacga	7440
cctcatgaaa	agcaaatagc	ctttcgtgag	caaacacgtt	atgtggcaag	actttcgccc	7500
ttagaccgcc	ccaatcctgg	agagaaaggg	acacaagagg	ctcttacctt	ccgtgatgat	7560
ggcagctatc	tgattgctgg	tggttttaggc	ggactggggg	tagtgggtgg	tcgttttctg	7620
gttacaaatg	gggctaata	ccttgtgcta	gtcggacgac	gtggtgcgag	ggaggaacag	7680
caagctcaat	taagcgaact	agagcaactc	ggagcttccg	tgaaagtgtt	acaagccgat	7740
attgctgatg	cagaacaact	agcccaagca	ctttcagcag	taacctaccc	accattacgg	7800
ggtgttatcc	atgcccaggg	tacattgaac	gatgggatcc	tacagcagca	aagttggcaa	7860
gccttttaaag	aagtgatgaa	tcccaaggta	gcaggtgcgt	ggaacctaca	tatactgaca	7920
aaaaatcagc	cttttagactt	ctttgtcctg	ttctcctccg	ccacctcttt	gttaggtaac	7980
gctggacaag	ccaattcacg	cgcgcgcaat	gctttccttg	atgggttagc	ctcctatcgt	8040
cgtcacttag	gactaccgag	cctctcgatt	aattggggga	catggagcga	agtgggaatt	8100
gcggtcgcag	ttgaactaga	taagttgtcc	agcaaacagg	gagagggaa	cattacgcta	8160
ggacagggct	tacaaattct	tgagcagttg	ctcaaagacg	agaatggggg	gtatcaagtg	8220
ggtgtcatgc	ctatcaagtg	gacacaattc	ttagcaaggc	aattgactcc	gcagccgttc	8280
ttcagcgatg	catgaagag	tattgacacc	ctgtaggtga	aactaacctt	gcaggagcgg	8340
gactcttgcc	cccaagggtta	cgggcataat	attcgagagc	aattagagaa	cgctccgccc	8400
aaagagggtc	tgactctctt	gcaggctcat	gttcgggagc	aggtttccca	agttttgggg	8460
atagacacga	agacattatt	ggcagaacaa	gacgtgggtt	tctttaccct	ggggatggat	8520
tcgctgacct	ctgtcgagtt	aagaacagg	ttacaagcca	gtttgggctg	ctctctttct	8580
tccacttttg	cttttgacta	tccaacacaa	caggctcttg	tgaattatct	tgccaatgaa	8640
ttgctgggaa	cccctgagca	gctacaagag	cctgaatctg	atgaagaaga	tcagatatcg	8700
tcaatggatg	acatcgtgca	gttgctgtcc	gcgaaactag	agatggaaat	ttaa	8754

<210> 100
 <211> 2917
 <212> PRT
 <213> *Cylindrospermopsis raciborskii* AWT205

5

<400> 100

```

Met Gln Lys Arg Glu Ser Pro Gln Ile Leu Phe Asp Gly Asn Gly Thr
1      5      10      15
Gln Ser Glu Phe Pro Asp Ser Cys Ile His His Leu Phe Glu Asp Gln
20      25      30
Ala Ala Lys Arg Pro Asp Ala Ile Ala Leu Ile Asp Gly Glu Gln Ser
35      40      45
Leu Thr Tyr Gly Glu Leu Asn Val Arg Ala Asn His Leu Ala Gln His
50      55      60
Leu Leu Ser Leu Gly Cys Gln Pro Asp Asp Leu Leu Ala Ile Cys Ile
65      70      75      80
Glu Arg Ser Ala Glu Leu Phe Ile Gly Leu Leu Gly Ile Leu Lys Ala
85      90      95
Gly Cys Ala Tyr Val Pro Leu Asp Val Gly Tyr Pro Gly Asp Arg Ile
100     105     110
Glu Tyr Met Leu Arg Asp Ser Asp Ala Arg Ile Leu Leu Thr Ser Thr
115     120     125
Asp Val Ala Lys Lys Leu Ala Leu Thr Ile Pro Ala Leu Gln Glu Cys
130     135     140
Gln Thr Val Tyr Leu Asp Gln Glu Ile Phe Glu Tyr Asp Phe His Phe
145     150     155     160
Leu Ala Ile Ala Lys Leu Leu His Asn Gln Tyr Leu Arg Leu Leu His
165     170     175
Phe Tyr Phe Tyr Thr Leu Ile Gln Gln Cys Gln Ala Thr Ser Val Ser
180     185     190
Gln Gly Ile Gln Thr Gln Val Leu Pro Asn Asn Leu Ala Tyr Cys Ile
195     200     205
Tyr Thr Ser Gly Ser Thr Gly Asn Pro Lys Gly Ile Leu Met Glu His
210     215     220
Arg Ser Leu Val Asn Met Leu Trp Trp His Gln Gln Thr Arg Pro Ser
225     230     235     240
Val Gln Gly Val Arg Thr Leu Gln Phe Cys Ala Val Ser Phe Asp Phe
245     250     255
Ser Cys His Glu Ile Phe Ser Thr Leu Cys Leu Gly Gly Ile Leu Val
260     265     270
Leu Val Pro Glu Ala Val Arg Gln Asn Pro Phe Ala Leu Ala Glu Phe
275     280     285
Ile Ser Gln Gln Lys Ile Glu Lys Leu Phe Leu Pro Val Ile Ala Leu
290     295     300
Leu Gln Leu Ala Glu Ala Val Asn Gly Asn Lys Ser Thr Ser Leu Ala
305     310     315     320

```

Leu Cys Glu Val Ile Thr Thr Gly Glu Gln Met Gln Ile Thr Pro Ala
 325 330 335
 Val Ala Asn Leu Phe Gln Lys Thr Gly Ala Met Leu His Asn His Tyr
 340 345 350
 Gly Ala Thr Glu Phe Gln Asp Ala Thr Thr His Thr Leu Lys Gly Asn
 355 360 365
 Pro Glu Gly Trp Pro Thr Leu Val Pro Val Gly Arg Pro Leu His Asn
 370 375 380
 Val Gln Val Tyr Ile Leu Asp Glu Ala Gln Gln Pro Val Pro Leu Gly
 385 390 395 400
 Gly Glu Gly Glu Phe Cys Ile Gly Gly Ile Gly Leu Ala Arg Gly Tyr
 405 410 415
 His Asn Leu Pro Asp Leu Thr Asn Glu Lys Phe Ile Pro Asn Pro Phe
 420 425 430
 Gly Ala Asn Glu Asn Ala Lys Lys Leu Tyr Arg Thr Gly Asp Leu Ala
 435 440 445
 Arg Tyr Leu Pro Asp Gly Thr Ile Glu His Leu Gly Arg Ile Asp His
 450 455 460
 Gln Val Lys Ile Arg Gly Phe Arg Val Glu Leu Gly Glu Ile Glu Ser
 465 470 475 480
 Val Leu Ala Ser His Gln Ala Val Arg Glu Cys Ala Val Val Ala Arg
 485 490 495
 Glu Ile Ala Gly His Thr Gln Leu Val Gly Tyr Ile Ile Ala Lys Asp
 500 505 510
 Thr Leu Asn Leu Ser Phe Asp Lys Leu Glu Pro Ile Leu Arg Gln Tyr
 515 520 525
 Ser Glu Ala Val Leu Pro Glu Tyr Met Ile Pro Thr Arg Phe Ile Asn
 530 535 540
 Ile Ser Asn Met Pro Leu Thr Pro Ser Gly Lys Leu Asp Arg Arg Ala
 545 550 555 560
 Leu Pro Asp Pro Lys Gly Asp Arg Pro Ala Leu Ser Thr Pro Leu Val
 565 570 575
 Lys Pro Arg Thr Gln Thr Glu Lys Arg Leu Ala Glu Ile Trp Gly Ser
 580 585 590
 Tyr Leu Ala Val Asp Ile Val Gly Thr His Asp Asn Phe Phe Asp Leu
 595 600 605
 Gly Gly Thr Ser Leu Leu Leu Thr Gln Ala His Lys Phe Leu Cys Glu
 610 615 620
 Thr Phe Asn Ile Asn Leu Ser Ala Val Ser Leu Phe Gln Tyr Pro Thr
 625 630 635 640
 Ile Gln Thr Leu Ala Gln Tyr Ile Asp Cys Gln Gly Asp Thr Thr Ser
 645 650 655
 Ser Asp Thr Ala Ser Arg His Lys Lys Val Arg Lys Lys Gln Ser Gly
 660 665 670
 Asp Ser Asn Asp Ile Ala Ile Ile Ser Val Ala Gly Arg Phe Pro Gly
 675 680 685
 Ala Glu Thr Ile Glu Gln Phe Trp His Asn Leu Cys Asn Gly Val Glu
 690 695 700
 Ser Ile Thr Leu Phe Ser Asp Asp Glu Leu Glu Gln Thr Leu Pro Glu
 705 710 715 720
 Leu Phe Asn Asn Pro Ala Tyr Val Lys Ala Gly Ala Val Leu Glu Gly
 725 730 735
 Val Glu Leu Phe Asp Ala Thr Phe Phe Gly Tyr Ser Pro Lys Glu Ala
 740 745 750
 Ala Val Thr Asp Pro Gln Gln Arg Ile Leu Leu Glu Cys Ala Trp Glu
 755 760 765
 Ala Phe Glu Arg Ala Gly Tyr Asn Pro Glu Thr Tyr Pro Glu Pro Val

770		775		780
Gly Val Tyr Ala Gly Ser Ser Leu Ser Thr Tyr Leu Leu Asn Asn Ile				
785		790		800
Gly Ser Ala Leu Gly Ile Ile Thr Glu Gln Pro Phe Ile Glu Thr Asp				
	805		810	815
Met Glu Gln Phe Gln Ala Lys Ile Gly Asn Asp Arg Ser Tyr Leu Ala				
	820		825	830
Thr Arg Ile Ser Tyr Lys Leu Asn Leu Lys Gly Pro Ser Val Asn Val				
	835		840	845
Gln Thr Ala Cys Ser Thr Ser Leu Val Ala Val His Met Ala Cys Gln				
	850		855	860
Ser Leu Ile Ser Gly Glu Cys Gln Met Ala Leu Ala Gly Gly Ile Ser				
865		870		880
Val Val Val Pro Gln Lys Gly Gly Tyr Leu Tyr Glu Glu Gly Met Val				
	885		890	895
Arg Ser Gln Asp Gly His Cys Arg Ala Phe Asp Ala Glu Ala Gln Gly				
	900		905	910
Thr Ile Phe Gly Asn Gly Gly Gly Leu Val Leu Leu Lys Arg Leu Gln				
	915		920	925
Asp Ala Leu Asp Asp Asn Asp Asn Ile Met Ala Val Ile Lys Ala Thr				
	930		935	940
Ala Ile Asn Asn Asp Gly Ala Leu Lys Met Gly Tyr Thr Ala Pro Ser				
945		950		960
Val Asp Gly Gln Ala Asp Val Ile Ser Glu Ala Ile Ala Ile Ala Asp				
	965		970	975
Ile Asp Ala Ser Thr Ile Gly Tyr Val Glu Ala His Gly Thr Ala Thr				
	980		985	990
Gln Leu Gly Asp Pro Ile Glu Val Ala Gly Leu Ala Arg Ala Phe Gln				
	995		1000	1005
Arg Ser Thr Asp Ser Val Leu Gly Lys Gln Gln Cys Ala Ile Gly				
	1010		1015	1020
Ser Val Lys Thr Asn Ile Gly His Leu Asp Glu Ala Ala Gly Ile				
	1025		1030	1035
Ala Gly Leu Ile Lys Ala Ala Leu Ala Leu Gln Tyr Gly Gln Ile				
	1040		1045	1050
Pro Pro Ser Leu His Tyr Ala Asn Pro Asn Pro Arg Ile Asp Phe				
	1055		1060	1065
Asp Ala Thr Pro Phe Phe Val Asn Thr Glu Leu Arg Glu Trp Ser				
	1070		1075	1080
Arg Asn Gly Tyr Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Val				
	1085		1090	1095
Gly Gly Thr Asn Ser His Ile Val Leu Glu Glu Ser Pro Val Lys				
	1100		1105	1110
Gln Pro Thr Leu Phe Ser Ser Leu Pro Glu Arg Ser His His Leu				
	1115		1120	1125
Leu Thr Leu Ser Ala His Thr Gln Glu Ala Leu His Glu Leu Val				
	1130		1135	1140
Gln Arg Tyr Ile Gln His Asn Glu Thr His Leu Asp Ile Asn Leu				
	1145		1150	1155
Gly Asp Leu Cys Phe Thr Ala Asn Thr Gly Arg Lys His Phe Glu				
	1160		1165	1170
His Arg Leu Ala Val Val Ala Glu Ser Ile Pro Gly Leu Gln Ala				
	1175		1180	1185
Gln Leu Glu Thr Ala Gln Thr Ala Ile Ser Ala Gln Lys Lys Asn				
	1190		1195	1200
Ala Pro Pro Thr Ile Ala Phe Leu Phe Thr Gly Gln Gly Ser Gln				
	1205		1210	1215

Tyr Ile	Asn Met Gly Arg Thr	Leu Tyr Asp Thr Glu	Ser Thr Phe
1220	1225	1230	
Arg Ala	Ala Leu Asp Arg Cys	Glu Thr Ile Leu Gln	Asn Leu Gly
1235	1240	1245	
Ile Glu	Ser Ile Leu Ser Val	Ile Phe Gly Ser Ser	Glu His Gly
1250	1255	1260	
Leu Ser	Leu Asp Asp Thr Ala	Tyr Thr Gln Pro Ala	Leu Phe Ala
1265	1270	1275	
Ile Glu	Tyr Ala Leu Tyr Gln	Leu Trp Lys Ser Trp	Gly Ile Gln
1280	1285	1290	
Pro Ser	Val Val Ile Gly His	Ser Val Gly Glu Tyr	Val Ser Ala
1295	1300	1305	
Cys Val	Ala Gly Val Phe Ser	Leu Glu Asp Gly Leu	Lys Leu Ile
1310	1315	1320	
Ala Glu	Arg Gly Arg Leu Ile	Gln Ala Leu Pro Arg	Asp Gly Ser
1325	1330	1335	
Met Val	Ser Val Met Ala Ser	Glu Lys Arg Ile Ala	Asp Ile Ile
1340	1345	1350	
Leu Pro	Tyr Gly Gly Gln Val	Gly Ile Ala Ala Ile	Asn Gly Pro
1355	1360	1365	
Gln Ser	Val Val Ile Ser Gly	Gln Gln Gln Ala Ile	Asp Ala Ile
1370	1375	1380	
Cys Ala	Ile Leu Glu Thr Glu	Gly Ile Lys Ser Lys	Lys Leu Asn
1385	1390	1395	
Val Ser	His Ala Phe His Ser	Pro Leu Val Glu Ala	Met Leu Asp
1400	1405	1410	
Ser Phe	Leu Gln Val Ala Gln	Glu Val Thr Tyr Ser	Gln Pro Gln
1415	1420	1425	
Ile Lys	Leu Ile Ser Asn Val	Thr Gly Thr Leu Ala	Ser His Glu
1430	1435	1440	
Ser Cys	Pro Asp Glu Leu Pro	Ile Thr Thr Ala Glu	Tyr Trp Val
1445	1450	1455	
Arg His	Val Arg Gln Pro Val	Arg Phe Ala Ala Gly	Met Glu Ser
1460	1465	1470	
Leu Glu	Gly Gln Gly Val Asn	Val Phe Ile Glu Ile	Gly Pro Lys
1475	1480	1485	
Pro Val	Leu Leu Gly Met Gly	Arg Asp Cys Leu Pro	Glu Gln Glu
1490	1495	1500	
Gly Leu	Trp Leu Pro Ser Leu	Arg Pro Lys Gln Asp	Asp Trp Gln
1505	1510	1515	
Gln Val	Leu Ser Ser Leu Arg	Asp Leu Tyr Leu Ala	Gly Val Thr
1520	1525	1530	
Val Asp	Trp Ser Ser Phe Asp	Gln Gly Tyr Ala Arg	Arg Arg Val
1535	1540	1545	
Pro Leu	Pro Thr Tyr Pro Trp	Gln Arg Glu Arg His	Trp Val Glu
1550	1555	1560	
Pro Ile	Ile Arg Gln Arg Gln	Ser Val Leu Gln Ala	Thr Asn Thr
1565	1570	1575	
Thr Lys	Leu Thr Arg Asn Ala	Ser Val Ala Gln His	Pro Leu Leu
1580	1585	1590	
Gly Gln	Arg Leu His Leu Ser	Arg Thr Gln Glu Ile	Tyr Phe Gln
1595	1600	1605	
Thr Phe	Ile His Ser Asp Phe	Pro Ile Trp Val Ala	Asp His Lys
1610	1615	1620	
Val Phe	Gly Asn Val Ile Ile	Pro Gly Val Ala Tyr	Phe Glu Met
1625	1630	1635	
Ala Leu	Ala Ala Gly Lys Ala	Leu Lys Pro Asp Ser	Ile Phe Trp

1640	1645	1650
Leu Glu Asp Val Ser Ile	Ala Gln Ala Leu Ile	Ile Pro Asp Glu
1655	1660	1665
Gly Gln Thr Val Gln Ile	Val Leu Ser Pro Gln	Glu Glu Ser Ala
1670	1675	1680
Tyr Phe Phe Glu Ile Leu	Ser Leu Glu Lys Glu	Asn Ser Trp Val
1685	1690	1695
Leu His Ala Ser Gly Lys	Leu Val Ala Gln Glu	Gln Val Leu Glu
1700	1705	1710
Thr Glu Pro Ile Asp Leu	Ile Ala Leu Gln Ala	His Cys Ser Glu
1715	1720	1725
Glu Val Ser Val Asp Val	Leu Tyr Gln Glu Glu	Met Ala Arg Arg
1730	1735	1740
Leu Asp Met Gly Pro Met	Met Arg Gly Val Lys	Gln Leu Trp Arg
1745	1750	1755
Tyr Pro Leu Ser Phe Ala	Lys Ser His Asp Ala	Ile Ala Leu Ala
1760	1765	1770
Lys Val Ser Leu Pro Glu	Ile Leu Leu His Glu	Ser Asn Ala Tyr
1775	1780	1785
Gln Phe His Pro Val Ile	Leu Asp Ala Gly Leu	Gln Met Ile Thr
1790	1795	1800
Val Ser Tyr Pro Glu Ala	Asn Gln Gly Gln Thr	Tyr Val Pro Val
1805	1810	1815
Gly Ile Glu Gly Leu Gln	Val Tyr Gly Arg Pro	Ser Ser Glu Leu
1820	1825	1830
Trp Cys Arg Ala Gln Tyr	Arg Pro Pro Leu Asp	Thr Asp Gln Arg
1835	1840	1845
Gln Gly Ile Asp Leu Leu	Pro Lys Lys Leu Ile	Ala Asp Leu His
1850	1855	1860
Leu Phe Asp Thr Gln Gly	Arg Val Val Ala Ile	Met Phe Gly Val
1865	1870	1875
Gln Ser Val Leu Val Gly	Arg Glu Ala Met Leu	Arg Ser Gln Asp
1880	1885	1890
Thr Trp Arg Asn Trp Leu	Tyr Gln Val Leu Trp	Lys Pro Gln Ala
1895	1900	1905
Cys Phe Gly Leu Leu Pro	Asn Tyr Leu Pro Thr	Pro Asp Lys Ile
1910	1915	1920
Arg Lys Arg Leu Glu Thr	Lys Leu Ala Thr Leu	Ile Ile Glu Ala
1925	1930	1935
Asn Leu Ala Thr Tyr Ala	Ile Ala Tyr Thr Gln	Leu Glu Arg Leu
1940	1945	1950
Ser Leu Ala Tyr Val Val	Ala Ala Phe Arg Gln	Met Gly Trp Leu
1955	1960	1965
Phe Gln Pro Gly Glu Arg	Phe Ser Thr Ala Gln	Lys Val Ser Ala
1970	1975	1980
Leu Gly Ile Val Asp Gln	His Arg Gln Leu Phe	Ala Arg Leu Leu
1985	1990	1995
Asp Ile Leu Ala Glu Ala	Asp Ile Leu Arg Ser	Glu Asn Leu Met
2000	2005	2010
Thr Ile Trp Glu Val Ile	Ser Tyr Pro Glu Thr	Ile Asp Ile Gln
2015	2020	2025
Val Leu Leu Asp Asp Leu	Glu Ala Lys Glu Ala	Glu Ala Glu Val
2030	2035	2040
Thr Leu Val Ser Arg Cys	Ser Ala Lys Leu Ala	Glu Val Leu Gln
2045	2050	2055
Gly Lys Cys Asp Pro Ile	Gln Leu Leu Phe Pro	Ala Gly Asp Thr
2060	2065	2070

Thr	Thr	Leu	Ser	Lys	Leu	Tyr	Arg	Glu	Ala	Pro	Val	Leu	Gly	Val
	2075					2080					2085			
Thr	Asn	Thr	Leu	Val	Gln	Glu	Ala	Leu	Leu	Ser	Ala	Leu	Glu	Gln
	2090					2095					2100			
Leu	Pro	Pro	Glu	Arg	Gly	Trp	Arg	Ile	Leu	Glu	Ile	Gly	Ala	Gly
	2105					2110					2115			
Thr	Gly	Gly	Thr	Thr	Ala	Tyr	Leu	Leu	Pro	His	Leu	Pro	Gly	Asp
	2120					2125					2130			
Gln	Thr	Lys	Tyr	Val	Phe	Thr	Asp	Ile	Ser	Ala	Phe	Phe	Leu	Ala
	2135					2140					2145			
Lys	Ala	Glu	Glu	Arg	Phe	Lys	Asp	Tyr	Pro	Phe	Val	Arg	Tyr	Gln
	2150					2155					2160			
Val	Leu	Asp	Ile	Glu	Gln	Ala	Pro	Gln	Ala	Gln	Gly	Phe	Glu	Pro
	2165					2170					2175			
Gln	Ile	Tyr	Asp	Leu	Ile	Val	Ala	Ala	Asp	Val	Leu	His	Ala	Thr
	2180					2185					2190			
Ser	Asp	Leu	Arg	Gln	Thr	Leu	Val	His	Ile	Arg	Gln	Leu	Leu	Ala
	2195					2200					2205			
Pro	Gly	Gly	Met	Leu	Ile	Leu	Met	Glu	Asp	Ser	Glu	Pro	Ala	Arg
	2210					2215					2220			
Trp	Ala	Asp	Leu	Thr	Phe	Gly	Leu	Thr	Glu	Gly	Trp	Trp	Lys	Phe
	2225					2230					2235			
Thr	Asp	His	Asp	Leu	Arg	Pro	Asn	His	Pro	Leu	Leu	Ser	Pro	Glu
	2240					2245					2250			
Gln	Trp	Gln	Ile	Leu	Leu	Ser	Glu	Met	Gly	Phe	Ser	Gln	Thr	Thr
	2255					2260					2265			
Ala	Leu	Trp	Pro	Lys	Ile	Asp	Ser	Pro	His	Lys	Leu	Pro	Arg	Glu
	2270					2275					2280			
Ala	Val	Ile	Val	Ala	Arg	Asn	Glu	Pro	Ala	Ile	Arg	Lys	Pro	Arg
	2285					2290					2295			
Arg	Trp	Leu	Ile	Leu	Ala	Asp	Glu	Glu	Ile	Gly	Gly	Leu	Leu	Ala
	2300					2305					2310			
Lys	Gln	Leu	Arg	Glu	Glu	Gly	Glu	Asp	Cys	Ile	Leu	Leu	Leu	Pro
	2315					2320					2325			
Gly	Glu	Lys	Tyr	Thr	Glu	Arg	Asp	Ser	Gln	Thr	Phe	Thr	Ile	Asn
	2330					2335					2340			
Pro	Gly	Asp	Ile	Glu	Glu	Trp	Gln	Gln	Leu	Leu	Asn	Arg	Val	Pro
	2345					2350					2355			
Asn	Ile	Gln	Glu	Ile	Val	His	Cys	Trp	Ser	Met	Val	Ser	Thr	Asp
	2360					2365					2370			
Leu	Asp	Arg	Ala	Thr	Ile	Phe	Ser	Cys	Ser	Ser	Thr	Leu	His	Leu
	2375					2380					2385			
Val	Gln	Ala	Leu	Ala	Asn	Tyr	Pro	Lys	Asn	Pro	Arg	Leu	Ser	Leu
	2390					2395					2400			
Val	Thr	Leu	Gly	Ala	Gln	Ala	Val	Asn	Glu	His	His	Val	Gln	Asn
	2405					2410					2415			
Val	Val	Gly	Ala	Ala	Leu	Trp	Gly	Met	Gly	Lys	Val	Ile	Ala	Leu
	2420					2425					2430			
Glu	His	Pro	Glu	Leu	Gln	Val	Ala	Gln	Met	Asp	Leu	Asp	Pro	Asn
	2435					2440					2445			
Gly	Lys	Val	Lys	Ala	Gln	Val	Glu	Val	Leu	Arg	Asp	Glu	Leu	Leu
	2450					2455					2460			
Ala	Arg	Lys	Asp	Pro	Ala	Ser	Ala	Met	Ser	Val	Pro	Asp	Leu	Gln
	2465					2470					2475			
Thr	Arg	Pro	His	Glu	Lys	Gln	Ile	Ala	Phe	Arg	Glu	Gln	Thr	Arg
	2480					2485					2490			
Tyr	Val	Ala	Arg	Leu	Ser	Pro	Leu	Asp	Arg	Pro	Asn	Pro	Gly	Glu

2495	Lys Gly Thr Gln Glu Ala	2500	Leu Thr Phe Arg Asp	2505	Asp Gly Ser Tyr
2510	Leu Ile Ala Gly Gly Leu	2515	Gly Gly Leu Gly Leu	2520	Val Val Ala Arg
2525	Phe Leu Val Thr Asn Gly	2530	Ala Lys Tyr Leu Val	2535	Leu Val Gly Arg
2540	Arg Gly Ala Arg Glu Glu	2545	Gln Gln Ala Gln Leu	2550	Ser Glu Leu Glu
2555	Gln Leu Gly Ala Ser Val	2560	Lys Val Leu Gln Ala	2565	Asp Ile Ala Asp
2570	Ala Glu Gln Leu Ala Gln	2575	Ala Leu Ser Ala Val	2580	Thr Tyr Pro Pro
2585	Leu Arg Gly Val Ile His	2590	Ala Ala Gly Thr Leu	2595	Asn Asp Gly Ile
2600	Leu Gln Gln Gln Ser Trp	2605	Gln Ala Phe Lys Glu	2610	Val Met Asn Pro
2615	Lys Val Ala Gly Ala Trp	2620	Asn Leu His Ile Leu	2625	Thr Lys Asn Gln
2630	Pro Leu Asp Phe Phe Val	2635	Leu Phe Ser Ser Ala	2640	Thr Ser Leu Leu
2645	Gly Asn Ala Gly Gln Ala	2650	Asn His Ala Ala Ala	2655	Asn Ala Phe Leu
2660	Asp Gly Leu Ala Ser Tyr	2665	Arg Arg His Leu Gly	2670	Leu Pro Ser Leu
2675	Ser Ile Asn Trp Gly Thr	2680	Trp Ser Glu Val Gly	2685	Ile Ala Ala Arg
2690	Leu Glu Leu Asp Lys Leu	2695	Ser Ser Lys Gln Gly	2700	Glu Gly Thr Ile
2705	Thr Leu Gly Gln Gly Leu	2710	Gln Ile Leu Glu Gln	2715	Leu Leu Lys Asp
2720	Glu Asn Gly Val Tyr Gln	2725	Val Gly Val Met Pro	2730	Ile Asn Trp Thr
2735	Gln Phe Leu Ala Arg Gln	2740	Leu Thr Pro Gln Pro	2745	Phe Phe Ser Asp
2750	Ala Met Lys Ser Ile Asp	2755	Thr Ser Val Gly Lys	2760	Leu Thr Leu Gln
2765	Glu Arg Asp Ser Cys Pro	2770	Gln Gly Tyr Gly His	2775	Asn Ile Arg Glu
2780	Gln Leu Glu Asn Ala Pro	2785	Pro Lys Glu Gly Leu	2790	Thr Leu Leu Gln
2795	Ala His Val Arg Glu Gln	2800	Val Ser Gln Val Leu	2805	Gly Ile Asp Thr
2810	Lys Thr Leu Leu Ala Glu	2815	Gln Asp Val Gly Phe	2820	Phe Thr Leu Gly
2825	Met Asp Ser Leu Thr Ser	2830	Val Glu Leu Arg Asn	2835	Arg Leu Gln Ala
2840	Ser Leu Gly Cys Ser Leu	2845	Ser Ser Thr Leu Ala	2850	Phe Asp Tyr Pro
2855	Thr Gln Gln Ala Leu Val	2860	Asn Tyr Leu Ala Asn	2865	Glu Leu Leu Gly
2870	Thr Pro Glu Gln Leu Gln	2875	Glu Pro Glu Ser Asp	2880	Glu Glu Asp Gln
2885	Ile Ser Ser Met Asp Asp	2890	Ile Val Gln Leu Leu	2895	Ser Ala Lys Leu
2900	Glu Met Glu Ile	2905		2910	
2915					

<210> 101

<211> 5667

<212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 101

atggatgaaa	aactaagaac	atacgaacga	ttaatcaagc	aatcctatca	caagatagag	60
gctctggaag	ctgaagttaa	caggttgaag	caaaccat	gtgaacctat	cgccatcgctc	120
ggcatgggct	gtcggtttcc	tggtgcgaat	agtccagaag	cgttttggca	gttggtgtgt	180
gatgggggtg	atgctattcg	tgagatacca	aaaaatcgat	gggttgttga	tgccctacata	240
gatgaaaatt	tggaccgcgc	agacaagaca	tcaatgcgat	ttggcggggt	tgtcgagcaa	300
cttgagaagt	ttgatgccca	attctttggc	atatcaccgc	gagaagcggg	ttctcttgac	360
cctcagcaac	gtttgttatt	agaagtaagt	tgggaagcac	tggaaaatgc	agcggtgata	420
ccaccttcgg	caacgggcgt	attcgtcggg	attagtaacc	ttgattatcg	tgaaacgctc	480
ttgaagcaag	gagcaattgg	tacttatttt	gcttcgggta	atgcccatag	cacagccagt	540
ggtcgcttgt	cttactttct	cggctcgaca	ggccccgtgc	tctcgataga	tacagcttgt	600
tcttcgctgt	tggtcgctgt	acatcagtc	ctgataagtc	tgcgtcagcg	agaatgtgac	660
ttagcggttg	ttgggggagt	ccatcggtcg	atagccccag	aggaaagtgt	ctcggttagca	720
aaagcccata	tgttatctcc	cgatggtcgt	tgcaaagtct	ttgatgcgtc	ggcaaacggg	780
tatgtccgag	ccgaaggatg	tggtcatgata	gtcctcaaac	gattatcgga	cgcgcaagct	840
gatgggggata	aaatcttggc	gttgattcgc	gggtcagcca	taaatcaaga	cggtcgcacg	900
agtggcttga	ccgttccaaa	tggtccccaa	caagccgacg	tgattcgcca	agccctcgcc	960
aatagtggca	taagaccaga	acaagttaac	tatgtagaag	ctcatggcac	agggacttcc	1020
ctaggagacc	cgattgaggt	cggcgcggtg	ggaacgatct	ttaatcaacg	ctcccaacct	1080
ttaattattg	gttcagttta	aacaaatatt	gggcactctag	aagcagcagc	agggattgct	1140
ggactgatta	aagtcgtcct	tgccatgcag	catggagaaa	ttccacctaa	tttacacttt	1200
caccagccca	atcctcgcat	taactgggat	aaattgccaa	tcaggatccc	cacagaacga	1260
acagcttgcc	ctactggcga	tcgcacgcga	gggataagtt	cttctgggctt	tagtggcact	1320
aattctcatg	tcgtgttaga	ggaagcccca	aaaatagagc	cgtctacttt	agagattcat	1380
tcaaagcagt	atgtttttac	cttatcagca	gcgacacctc	aagcactaca	agaacttact	1440
cagcggttatg	taacttatct	cactgaacac	ttacaagaga	gtctggcgga	tatttgcttt	1500
acagccaaca	cagggcgcaa	acactttaga	catcgctttg	cagtagtagc	agagtctaaa	1560
acccagttgc	gccacaatt	ggaaacgttt	gccccaatcg	gagaggggca	ggggaagagg	1620
acatctctct	caaaaatagc	tttctctctt	acaggtcaag	gctcacagta	tgtggggatg	1680
gggcaagaac	tttatgagag	ccaaccaccc	ttccggcaaa	ccattgaccg	atgtgatgag	1740
attcttcggt	cactgttggg	caaatcaatc	ctctcaatac	tctatcccag	ccaacaaatg	1800
ggattggaaa	cgccatccca	aattgatgaa	accgcctata	ctcaacccac	tcttttttct	1860
cttgaatatg	cactggcgca	gttgtggcgc	tcctggggta	ttgagcctga	tgtgggtgatg	1920
gggcatagtg	tgggagaata	tgtggccgct	tgtgtggcgg	gtgtcttttc	tttagaggat	1980
ggactcaaac	taattgctga	aagaggccgt	ctgatgcaag	aattgcctcc	cgatggggcg	2040
atggtttcag	ttatggccaa	taaatcgcg	atagagcaag	caattcaatc	tgtcagccga	2100
gaggtttcta	ttgcggccat	caatggacct	gagagtgtgg	ttatctctgg	taaaaggag	2160
atattacaac	agattaccga	acatctggtt	gccgaaggca	ttaagacacg	ccaactgaag	2220
gtctctcatg	cctttcactc	accattgatg	gagccaatat	taggtcagtt	ccgccagtt	2280
gccaatacca	tcacctatcg	gccaccgcaa	attaaccttg	tctcaaatgt	cacaggcgga	2340
caggtgtata	aagaaatcgc	tactccgat	tattgggtga	gacatctgca	agagactgtc	2400
cgttttgcgg	atgggggttaa	ggtgttacat	gaacagaatg	tcaatttcat	gctcgaaatt	2460
gggtcccaaac	ccacactgct	gggcatgggt	gagttacaaa	gttctgagaa	tccattttct	2520
atgccaatga	tgatgcccg	tttgcgtcag	aatcgtagcg	actggcagca	gatgttgag	2580
agcttgagtc	aactctatgt	tcattggtgt	gagattgact	ggatcggttt	taataaagac	2640
tatgtgcgac	ataaagtgtg	cctgccgaca	tacccatggc	agaaggagcg	ttactgggta	2700
gaattggatc	aacagaagca	cgccgctaaa	aatctacatc	ctctactgga	caggtgcatg	2760
aagctgcctc	gtcataacga	aacaattttt	gagaaagaat	ttagtctaga	gacattgccc	2820
tttcttgctg	actatcgcat	ttatgggttca	gttgtgtcgc	caggtgcaag	ttatctatca	2880
atgatactaa	gtattgccga	gtcgtatgca	aatgggtcatt	tgaatggagg	gaatagtgca	2940
aagcaaacca	cttattttact	aaaggatgtc	acattcccag	tacctcttgt	gatctctgat	3000
gaggcaaat	acatggtgca	agttgcttgt	tctctctctt	gtgctgcgc	acacaatcgt	3060

```

ggcgacgaga cgcagtttga attgttcagt tttgctgaga atgtacctga aagtagcagt 3120
ataaatgctg atttttcagac acccattatt catgcaaaaag ggcaatttaa gcttgaagat 3180
acagcacctc ctaaagtgga gctagaagaa ctacaagcgg gttgtcccca agaaattgat 3240
ctcaaccttt tctatcaaac attcacagac aaagggttttg tttttggatc tcgttttcgc 3300
tggttagaac aaatctgggt gggcgatgga gaagcattgg cgcgtctgcg acaaccggaa 3360
agtattgaat cgttttaaagg atatgtgatt catcccgggt tgttggatgc ctgtacacaa 3420
gtcccatttg caatttcgct tgacgatgaa aataggcaat cagaaacgac aatgcccttt 3480
gcgctgaatg aattacgttg ttatcagcct gcaaacggac aaatgtgggt ggttcattgca 3540
acagaaaaag atagatatat atgggatgtt tctctgtttg atgagagcgg gcaagttatt 3600
gcggaattta taggtttaga agttcgtgct gctatgcccg aaggcttact aagggcagac 3660
ttttggcata actggctcta tacagtgaat tggcgatcgc aacctctaca aatcccagag 3720
gtgctggata ttaataagac aggtgcagaa acatggcttc tttttgcaca accagaggga 3780
ataggagcgg acttagccga atatttgag agccaaggaa agcactgtgt ttttgtagt 3840
cctgggagtg agtatacagt gaccgagcaa cacattggac gcaactggaca tcttgatgtg 3900
acgaaaactga caaaaattgt cacgatcaat cctgcttctc ctcatgacta taaatatttt 3960
ttagaaactc tgacggacat tagattacct tgtgaacata tactctattt atggaatcgt 4020
tatgatttaa caaatacttc taatcatcgg acagaattga ctgtaccaga tatagtctta 4080
aacttatgta ctagtcttac ttatttggta caagccctta gccacatggg tttttccccc 4140
aaattatggc taattacaca aaatagtcaa gcggttggtg gtgacttagc gaatttagaa 4200
atcgaaacaat cccattatg ggcattgggt cgaagcatcc gcgcggaaca ccctgaattt 4260
gattgccgtt gtttagattt tgacacgctc tcaaatatcg caccactctt gttgaaagag 4320
atgcaagcta tagactatga atctcaaat gcttaccgac aaggaaacgcg ctatgttgca 4380
cgactaatc gtaatcaatc agaattgtcac gcaccgatcc aaacaggaat ccgtcctgat 4440
ggcagctatt tgattacagg tggattaggg ggtctaggat tgcaggtagc actcgccctt 4500
gcggaacgctg gagcaagaca cttgatccct aatagtcgcc gtggtacggg ctccaaagaa 4560
gcccagttaa ttattgaccg actacgcca gaggatgta ggggtgattt gattgcggca 4620
gatgtctctg atgcggcaga tagcgaacga ctcttagtag aaagtcagcg caagacctct 4680
cttcgagggg ttgtccatgt tgcgggagtc ttggatgatg gcatcctgct ccaacaaaat 4740
caagagcgtt ttgaaaaagt gatggcgggt aaggtagcgc gagcttgga tctggacca 4800
cagagccaaa cctcgtatt agatttcttt gttgcgttct catctgttgc gtcgctcata 4860
gaagaaccag gacaagccaa ttacgccga gcgaatgct ttttggattc attaattgat 4920
tatcgtcaca taaagggatc taatagcttg agtatcaact ggggggcttg ggcagaagtc 4980
ggcatggcag ccaatttate atgggaacaa cggggaatcg cggcaatttc tccaaagcaa 5040
gggaggcata ttctcgtcca acttattcaa aaacttaatc agcatacaat cccccaagtt 5100
gctgtacaac cgaccaattg ggctgaatat ctatcccatg atggcgtgaa tatgccattc 5160
tatgaatatt ttacacacca cttgcgtaac gaaaaagaag ccaaattgcg gcaaacagca 5220
ggcagcacct cagaggaagt cagtctgcgg caacagcttc aaacactctc agagaaagac 5280
cgggatgccc ttttgatgga acatcttcaa aaaactgcga tcagagttct cggtttggca 5340
tctaatacaa aaattgatcc ctatcaggga ttgatgaata tgggactaga ctctttgatg 5400
gcggttgaat ttcggaatca cttgatacgt agtttagaac gccctctgcc agccactctg 5460
ctctttaatt gcccaacact tgattcattg catgattacc tagtcgcaaa aatgtttgat 5520
gatgcccctc agaaggcaga gcaaatggca caaccaacaa cactgacagc acacagcata 5580
tcaatagaat ccaaaataga tgataacgaa agcgtggatg acattgcaca aatgctggca 5640
caagcactca atatcgcctt tgagtag 5667

```

<210> 102

<211> 1888

5 <212> PRT

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 102

```

Met Asp Glu Lys Leu Arg Thr Tyr Glu Arg Leu Ile Lys Gln Ser Tyr
1          5          10          15
His Lys Ile Glu Ala Leu Glu Ala Glu Val Asn Arg Leu Lys Gln Thr
20          25          30
Gln Cys Glu Pro Ile Ala Ile Val Gly Met Gly Cys Arg Phe Pro Gly
35          40          45
10 Ala Asn Ser Pro Glu Ala Phe Trp Gln Leu Leu Cys Asp Gly Val Asp

```

50		55		60	
Ala Ile Arg Glu Ile	Pro Lys Asn Arg Trp	Val Val Asp Ala Tyr Ile			
65	70	75	80		
Asp Glu Asn Leu Asp	Arg Ala Asp Lys Thr	Ser Met Arg Phe Gly Gly			
	85	90	95		
Phe Val Glu Gln Leu	Glu Lys Phe Asp Ala Gln	Phe Phe Gly Ile Ser			
	100	105	110		
Pro Arg Glu Ala Val	Ser Leu Asp Pro Gln Gln	Arg Leu Leu Leu Glu			
	115	120	125		
Val Ser Trp Glu Ala	Leu Glu Asn Ala Ala	Val Ile Pro Pro Ser Ala			
	130	135	140		
Thr Gly Val Phe Val	Gly Ile Ser Asn Leu Asp	Tyr Arg Glu Thr Leu			
145	150	155	160		
Leu Lys Gln Gly Ala	Ile Gly Thr Tyr Phe Ala	Ser Gly Asn Ala His			
	165	170	175		
Ser Thr Ala Ser Gly	Arg Leu Ser Tyr Phe	Leu Gly Leu Thr Gly Pro			
	180	185	190		
Cys Leu Ser Ile Asp	Thr Ala Cys Ser Ser	Ser Leu Val Ala Val His			
	195	200	205		
Gln Ser Leu Ile Ser	Leu Arg Gln Arg Glu	Cys Asp Leu Ala Leu Val			
	210	215	220		
Gly Gly Val His Arg	Leu Ile Ala Pro Glu	Glu Ser Val Ser Leu Ala			
225	230	235	240		
Lys Ala His Met Leu	Ser Pro Asp Gly Arg	Cys Lys Val Phe Asp Ala			
	245	250	255		
Ser Ala Asn Gly Tyr	Val Arg Ala Glu Gly	Cys Gly Met Ile Val Leu			
	260	265	270		
Lys Arg Leu Ser Asp	Ala Gln Ala Asp Gly	Asp Lys Ile Leu Ala Leu			
	275	280	285		
Ile Arg Gly Ser Ala	Ile Asn Gln Asp Gly	Arg Thr Ser Gly Leu Thr			
	290	295	300		
Val Pro Asn Gly Pro	Gln Gln Ala Asp Val	Ile Arg Gln Ala Leu Ala			
305	310	315	320		
Asn Ser Gly Ile Arg	Pro Glu Gln Val Asn	Tyr Val Glu Ala His Gly			
	325	330	335		
Thr Gly Thr Ser Leu	Gly Asp Pro Ile Glu	Val Gly Ala Leu Gly Thr			
	340	345	350		
Ile Phe Asn Gln Arg	Ser Gln Pro Leu Ile	Ile Gly Ser Val Lys Thr			
	355	360	365		
Asn Ile Gly His Leu	Glu Ala Ala Ala Gly	Ile Ala Gly Leu Ile Lys			
	370	375	380		
Val Val Leu Ala Met	Gln His Gly Glu Ile	Pro Pro Asn Leu His Phe			
385	390	395	400		
His Gln Pro Asn Pro	Arg Ile Asn Trp Asp	Lys Leu Pro Ile Arg Ile			
	405	410	415		
Pro Thr Glu Arg Thr	Ala Trp Pro Thr Gly	Asp Arg Ile Ala Gly Ile			
	420	425	430		
Ser Ser Phe Gly Phe	Ser Gly Thr Asn Ser	His Val Val Leu Glu Glu			
	435	440	445		
Ala Pro Lys Ile Glu	Pro Ser Thr Leu Glu	Ile His Ser Lys Gln Tyr			
	450	455	460		
Val Phe Thr Leu Ser	Ala Ala Thr Pro Gln	Ala Leu Gln Glu Leu Thr			
465	470	475	480		
Gln Arg Tyr Val Thr	Tyr Leu Thr Glu His	Leu Gln Glu Ser Leu Ala			
	485	490	495		
Asp Ile Cys Phe Thr	Ala Asn Thr Gly Arg	Lys His Phe Arg His Arg			
	500	505	510		

Phe	Ala	Val	Val	Ala	Glu	Ser	Lys	Thr	Gln	Leu	Arg	Gln	Gln	Leu	Glu
		515					520					525			
Thr	Phe	Ala	Gln	Ser	Gly	Glu	Gly	Gln	Gly	Lys	Arg	Thr	Ser	Leu	Ser
		530				535					540				
Lys	Ile	Ala	Phe	Leu	Phe	Thr	Gly	Gln	Gly	Ser	Gln	Tyr	Val	Gly	Met
545					550					555					560
Gly	Gln	Glu	Leu	Tyr	Glu	Ser	Gln	Pro	Thr	Phe	Arg	Gln	Thr	Ile	Asp
				565					570					575	
Arg	Cys	Asp	Glu	Ile	Leu	Arg	Ser	Leu	Leu	Gly	Lys	Ser	Ile	Leu	Ser
			580					585					590		
Ile	Leu	Tyr	Pro	Ser	Gln	Gln	Met	Gly	Leu	Glu	Thr	Pro	Ser	Gln	Ile
		595					600					605			
Asp	Glu	Thr	Ala	Tyr	Thr	Gln	Pro	Thr	Leu	Phe	Ser	Leu	Glu	Tyr	Ala
		610				615					620				
Leu	Ala	Gln	Leu	Trp	Arg	Ser	Trp	Gly	Ile	Glu	Pro	Asp	Val	Val	Met
625					630					635					640
Gly	His	Ser	Val	Gly	Glu	Tyr	Val	Ala	Ala	Cys	Val	Ala	Gly	Val	Phe
				645					650					655	
Ser	Leu	Glu	Asp	Gly	Leu	Lys	Leu	Ile	Ala	Glu	Arg	Gly	Arg	Leu	Met
			660				665						670		
Gln	Glu	Leu	Pro	Pro	Asp	Gly	Ala	Met	Val	Ser	Val	Met	Ala	Asn	Lys
		675					680					685			
Ser	Arg	Ile	Glu	Gln	Ala	Ile	Gln	Ser	Val	Ser	Arg	Glu	Val	Ser	Ile
		690				695					700				
Ala	Ala	Ile	Asn	Gly	Pro	Glu	Ser	Val	Val	Ile	Ser	Gly	Lys	Arg	Glu
705					710					715					720
Ile	Leu	Gln	Gln	Ile	Thr	Glu	His	Leu	Val	Ala	Glu	Gly	Ile	Lys	Thr
				725					730					735	
Arg	Gln	Leu	Lys	Val	Ser	His	Ala	Phe	His	Ser	Pro	Leu	Met	Glu	Pro
			740					745					750		
Ile	Leu	Gly	Gln	Phe	Arg	Arg	Val	Ala	Asn	Thr	Ile	Thr	Tyr	Arg	Pro
		755					760					765			
Pro	Gln	Ile	Asn	Leu	Val	Ser	Asn	Val	Thr	Gly	Gly	Gln	Val	Tyr	Lys
		770				775					780				
Glu	Ile	Ala	Thr	Pro	Asp	Tyr	Trp	Val	Arg	His	Leu	Gln	Glu	Thr	Val
785					790				795						800
Arg	Phe	Ala	Asp	Gly	Val	Lys	Val	Leu	His	Glu	Gln	Asn	Val	Asn	Phe
				805					810					815	
Met	Leu	Glu	Ile	Gly	Pro	Lys	Pro	Thr	Leu	Leu	Gly	Met	Val	Glu	Leu
			820					825					830		
Gln	Ser	Ser	Glu	Asn	Pro	Phe	Ser	Met	Pro	Met	Met	Met	Pro	Ser	Leu
		835					840					845			
Arg	Gln	Asn	Arg	Ser	Asp	Trp	Gln	Gln	Met	Leu	Glu	Ser	Leu	Ser	Gln
					855					860					
Leu	Tyr	Val	His	Gly	Val	Glu	Ile	Asp	Trp	Ile	Gly	Phe	Asn	Lys	Asp
865					870					875					880
Tyr	Val	Arg	His	Lys	Val	Val	Leu	Pro	Thr	Tyr	Pro	Trp	Gln	Lys	Glu
				885					890					895	
Arg	Tyr	Trp	Val	Glu	Leu	Asp	Gln	Gln	Lys	His	Ala	Ala	Lys	Asn	Leu
			900					905					910		
His	Pro	Leu	Leu	Asp	Arg	Cys	Met	Lys	Leu	Pro	Arg	His	Asn	Glu	Thr
		915					920					925			
Ile	Phe	Glu	Lys	Glu	Phe	Ser	Leu	Glu	Thr	Leu	Pro	Phe	Leu	Ala	Asp
		930				935					940				
Tyr	Arg	Ile	Tyr	Gly	Ser	Val	Val	Ser	Pro	Gly	Ala	Ser	Tyr	Leu	Ser
945					950					955					960
Met	Ile	Leu	Ser	Ile	Ala	Glu	Ser	Tyr	Ala	Asn	Gly	His	Leu	Asn	Gly

965										970					975				
Gly	Asn	Ser	Ala	Lys	Gln	Thr	Thr	Tyr	Leu	Leu	Lys	Asp	Val	Thr	Phe				
980										985					990				
Pro	Val	Pro	Leu	Val	Ile	Ser	Asp	Glu	Ala	Asn	Tyr	Met	Val	Gln	Val				
995										1000					1005				
Ala	Cys	Ser	Leu	Ser	Cys	Ala	Ala	Pro	His	Asn	Arg	Gly	Asp	Glu					
1010						1015					1020								
Thr	Gln	Phe	Glu	Leu	Phe	Ser	Phe	Ala	Glu	Asn	Val	Pro	Glu	Ser					
1025						1030					1035								
Ser	Ser	Ile	Asn	Ala	Asp	Phe	Gln	Thr	Pro	Ile	Ile	His	Ala	Lys					
1040						1045					1050								
Gly	Gln	Phe	Lys	Leu	Glu	Asp	Thr	Ala	Pro	Pro	Lys	Val	Glu	Leu					
1055						1060					1065								
Glu	Glu	Leu	Gln	Ala	Gly	Cys	Pro	Gln	Glu	Ile	Asp	Leu	Asn	Leu					
1070						1075					1080								
Phe	Tyr	Gln	Thr	Phe	Thr	Asp	Lys	Gly	Phe	Val	Phe	Gly	Ser	Arg					
1085						1090					1095								
Phe	Arg	Trp	Leu	Glu	Gln	Ile	Trp	Val	Gly	Asp	Gly	Glu	Ala	Leu					
1100						1105					1110								
Ala	Arg	Leu	Arg	Gln	Pro	Glu	Ser	Ile	Glu	Ser	Phe	Lys	Gly	Tyr					
1115						1120					1125								
Val	Ile	His	Pro	Gly	Leu	Leu	Asp	Ala	Cys	Thr	Gln	Val	Pro	Phe					
1130						1135					1140								
Ala	Ile	Ser	Ser	Asp	Asp	Glu	Asn	Arg	Gln	Ser	Glu	Thr	Thr	Met					
1145						1150					1155								
Pro	Phe	Ala	Leu	Asn	Glu	Leu	Arg	Cys	Tyr	Gln	Pro	Ala	Asn	Gly					
1160						1165					1170								
Gln	Met	Trp	Trp	Val	His	Ala	Thr	Glu	Lys	Asp	Arg	Tyr	Thr	Trp					
1175						1180					1185								
Asp	Val	Ser	Leu	Phe	Asp	Glu	Ser	Gly	Gln	Val	Ile	Ala	Glu	Phe					
1190						1195					1200								
Ile	Gly	Leu	Glu	Val	Arg	Ala	Ala	Met	Pro	Glu	Gly	Leu	Leu	Arg					
1205						1210					1215								
Ala	Asp	Phe	Trp	His	Asn	Trp	Leu	Tyr	Thr	Val	Asn	Trp	Arg	Ser					
1220						1225					1230								
Gln	Pro	Leu	Gln	Ile	Pro	Glu	Val	Leu	Asp	Ile	Asn	Lys	Thr	Gly					
1235						1240					1245								
Ala	Glu	Thr	Trp	Leu	Leu	Phe	Ala	Gln	Pro	Glu	Gly	Ile	Gly	Ala					
1250						1255					1260								
Asp	Leu	Ala	Glu	Tyr	Leu	Gln	Ser	Gln	Gly	Lys	His	Cys	Val	Phe					
1265						1270					1275								
Val	Val	Pro	Gly	Ser	Glu	Tyr	Thr	Val	Thr	Glu	Gln	His	Ile	Gly					
1280						1285					1290								
Arg	Thr	Gly	His	Leu	Asp	Val	Thr	Lys											

Leu Glu	Ile Glu	Gln Ser	Pro	Leu Trp	Ala Leu	Gly	Arg Ser	Ile	
1400			1405			1410			
Arg Ala	Glu His	Pro Glu	Phe	Asp Cys	Arg Cys	Leu	Asp Phe	Asp	
1415			1420			1425			
Thr Leu	Ser Asn	Ile Ala	Pro	Leu Leu	Leu Lys	Glu	Met Gln	Ala	
1430			1435			1440			
Ile Asp	Tyr Glu	Ser Gln	Ile	Ala Tyr	Arg Gln	Gly	Thr Arg	Tyr	
1445			1450			1455			
Val Ala	Arg Leu	Ile Arg	Asn	Gln Ser	Glu Cys	His	Ala Pro	Ile	
1460			1465			1470			
Gln Thr	Gly Ile	Arg Pro	Asp	Gly Ser	Tyr Leu	Ile	Thr Gly	Gly	
1475			1480			1485			
Leu Gly	Gly Leu	Gly Leu	Gln	Val Ala	Leu Ala	Leu	Ala Asp	Ala	
1490			1495			1500			
Gly Ala	Arg His	Leu Ile	Leu	Asn Ser	Arg Arg	Gly	Thr Val	Ser	
1505			1510			1515			
Lys Glu	Ala Gln	Leu Ile	Ile	Asp Arg	Leu Arg	Gln	Glu Asp	Val	
1520			1525			1530			
Arg Val	Asp Leu	Ile Ala	Ala	Asp Val	Ser Asp	Ala	Ala Asp	Ser	
1535			1540			1545			
Glu Arg	Leu Leu	Val Glu	Ser	Gln Arg	Lys Thr	Ser	Leu Arg	Gly	
1550			1555			1560			
Ile Val	His Val	Ala Gly	Val	Leu Asp	Asp Gly	Ile	Leu Leu	Gln	
1565			1570			1575			
Gln Asn	Gln Glu	Arg Phe	Glu	Lys Val	Met Ala	Ala	Lys Val	Arg	
1580			1585			1590			
Gly Ala	Trp His	Leu Asp	Gln	Gln Ser	Gln Thr	Leu	Asp Leu	Asp	
1595			1600			1605			
Phe Phe	Val Ala	Phe Ser	Ser	Val Ala	Ser Leu	Ile	Glu Glu	Pro	
1610			1615			1620			
Gly Gln	Ala Asn	Tyr Ala	Ala	Ala Asn	Ala Phe	Leu	Asp Ser	Leu	
1625			1630			1635			
Met Tyr	Tyr Arg	His Ile	Lys	Gly Ser	Asn Ser	Leu	Ser Ile	Asn	
1640			1645			1650			
Trp Gly	Ala Trp	Ala Glu	Val	Gly Met	Ala Ala	Asn	Leu Ser	Trp	
1655			1660			1665			
Glu Gln	Arg Gly	Ile Ala	Ala	Ile Ser	Pro Lys	Gln	Gly Arg	His	
1670			1675			1680			
Ile Leu	Val Gln	Leu Ile	Gln	Lys Leu	Asn Gln	His	Thr Ile	Pro	
1685			1690			1695			
Gln Val	Ala Val	Gln Pro	Thr	Asn Trp	Ala Glu	Tyr	Leu Ser	His	
1700			1705			1710			
Asp Gly	Val Asn	Met Pro	Phe	Tyr Glu	Tyr Phe	Thr	His His	Leu	
1715			1720			1725			
Arg Asn	Glu Lys	Glu Ala	Lys	Leu Arg	Gln Thr	Ala	Gly Ser	Thr	
1730			1735			1740			
Ser Glu	Glu Val	Ser Leu	Arg	Gln Gln	Leu Gln	Thr	Leu Ser	Glu	
1745			1750			1755			
Lys Asp	Arg Asp	Ala Leu	Leu	Met Glu	His Leu	Gln	Lys Thr	Ala	
1760			1765			1770			
Ile Arg	Val Leu	Gly Leu	Ala	Ser Asn	Gln Lys	Ile	Asp Pro	Tyr	
1775			1780			1785			
Gln Gly	Leu Met	Asn Met	Gly	Leu Asp	Ser Leu	Met	Ala Val	Glu	
1790			1795			1800			
Phe Arg	Asn His	Leu Ile	Arg	Ser Leu	Glu Arg	Pro	Leu Pro	Ala	
1805			1810			1815			
Thr Leu	Leu Phe	Asn Cys	Pro	Thr Leu	Asp Ser	Leu	His Asp	Tyr	
1820			1825			1830			
Leu Val	Ala Lys	Met Phe	Asp	Asp Ala	Pro Gln	Lys	Ala Glu	Gln	
1835			1840			1845			
Met Ala	Gln Pro	Thr Thr	Leu	Thr Ala	His Ser	Ile	Ser Ile	Glu	
1850			1855			1860			
Ser Lys	Ile Asp	Asp Asn	Glu	Ser Val	Asp Asp	Ile	Ala Gln	Met	
1865			1870			1875			
Leu Ala	Gln Ala	Leu Asn	Ile	Ala Phe	Glu				
1880			1885						

<211> 5004
 <212> ADN
 <213> *Cylindrospermopsis raciborskii* AWT205

5 <400> 103

atgagtcagc	ccaattatgg	catttttgatg	aaaaatgcgt	tgaacgaaat	aaatagccta	60
cgatcgcaac	tagctgcggt	agaagcccaa	aaaaatgagt	ctattgccat	tgttggtatg	120
agttgccgtt	ttccaggcgg	tgcaactact	ccagagcggt	tttgggtatt	actgcgcgag	180
ggtatatcag	ccattacaga	aatccctgct	gatcgctggg	atgttgataa	atattatgat	240
gctgacccca	catcgctcgg	taaaatgcat	actcgttacg	gcggttttct	gaatgaagtt	300
gatacatttg	agccatcatt	ctttaatat	gctgcccgtg	aagccggttag	catggatcca	360
cagcaacgct	tgctacttga	agtcagttgg	gaagctctgg	aatccggtaa	tattgttccct	420
gcaactcttt	ttgatagtcc	cactgggtga	tttatcggtg	ttgggtggtag	caactacaaa	480
tctttaatga	tcgaaaacag	gagtcggatc	gggaaaaccg	atttgtatga	gttaagtggc	540
actgatgtga	gtgttgctgc	cggcaggata	tcctatgtcc	tgggtttgat	gggtcccagt	600
tttgtgattg	atacagcttg	ttcatcttct	ttggctctcag	ttcatcaagc	ctgtcagagt	660
ctgctgcaga	gagaatgtga	tctagcacta	gctgggtggg	tcggtttact	cattgatcca	720
gatgagatga	ttggtctttc	tcaagggggg	atgctggcac	ctgatggtag	ttgtaaaaca	780
tttgatgcca	atgcaaatgg	ctatgtgcga	ggcgaaaggt	gtgggatgat	tgttctaaaa	840
cgtctctcgg	atgcaacagc	cgatggggat	aatattcttg	ccatcattcg	tgggtctatg	900
gttaaatcatg	atggtcatag	cagtgggtta	actgctccaa	gaggccccgc	acaagtctct	960
gtcattaagc	aagccttaga	tagagcaggt	attgcaccgg	atgccgtaag	ttatttagaa	1020
gcccattgta	caggcacacc	ccttggtgat	cctatcgaga	tggattcatt	gaacgaagtg	1080
tttggtcgga	gaacagaacc	actttgggtc	ggctcagtta	agacaaatat	tggctattta	1140
gaagccgcgt	ccggtattgc	agggctgatt	aagggtgtct	tgatgctaaa	aaacaagcag	1200
attcctcctc	acttgcattt	caagacacca	aatccatata	ttgattggaa	aaatctcccg	1260
gtcgaaatte	cgaccaccct	tcatgcttgg	gatgacaaga	cattgaagga	cagaaagcga	1320
attgcagggg	ttagtctctt	tagtttcagt	ggtactaacg	cccacattgt	attatctgaa	1380
gccccatcta	gcgaactaat	tagtaatcat	gcggcagtg	aaagaccatg	gcacttggtta	1440
acccttagtg	ctaagaatga	ggaagcggtg	gctaacttgg	ttgggcttta	tcagtcattt	1500
atttctacta	ctgatgcaag	tcttgccgat	atatgctaca	ctgctaatac	ggcacgaacc	1560
cattttttctc	atcgcccttg	tctatcggtc	acttcacaca	tccaaataga	ggctctttta	1620
gccgcttata	aggaaggggtc	ggtgagtttg	agcatcaatc	aagggtgtgt	cctttccaac	1680
agtcgtgcgc	cgaaggtcgc	ttttctcttt	acaggtcaag	gttcgcaata	tgtgcaaatg	1740
gctggagaac	tttatgagac	ccagcctact	ttccgtaatt	gcttagatcg	ctgtgccgaa	1800
atccttgcaat	ccatcttttc	atcgagaaac	agcccttggg	gaaacccact	gctttcggta	1860
ttatatccaa	accatgagtc	aaaggaaatt	gaccagacgg	cttataccca	acctgccctt	1920
tttgctgtag	aatatgccct	agcacagatg	tggcggtcgt	ggggaatcga	gccagatata	1980
gtaatgggtc	atagcatagg	tgaatatgtg	gcagcttggt	tggcggggat	cttttctctg	2040
gaggatgggtc	tcaaaacttg	tgccgaaaga	ggccgtttga	tgcaggcgct	accacaaaat	2100
ggcgagatgg	ttgttatatc	ggcctccctt	gaggaagtta	agccggctat	tcaatctgac	2160
cagcgagttg	tgatagcggc	ggtaaatgga	ccacgaagtg	tcgtcatttc	gggcgatcgc	2220
caagctgtgc	aagctttcac	caacacccta	gaagatcaag	gaatccgggtg	caagagactg	2280
tctgtttcac	acgctttcca	ctctccattg	atgaaaccaa	tggagcagga	gttcgcacag	2340
gtggccaggg	aaatcaacta	tagtctcca	aaaatagctc	ttgtcagtaa	tctaaccggc	2400
gacttgatgt	cacctgagtc	ttccctggag	gaaggagtga	tcgcttcccc	tggttactgg	2460

```

gtaaatcatt tatgcaatcc tgtcttgttc gctgatggta ttgcaactat gcaagcgag 2520
gatgtccaag tcttccttga agttggacca aaaccgacct tatcaggact agtgaacaa 2580
tattttgacg aggttgccca tagcgatcgc cctgtcacca ttcccacctt gcgccccaaag 2640
caacccaact ggcagacact attggagagt ttgggacaac tgtatgcgct tgggtgtccag 2700
gtaaattggg cgggctttga tagagattac accagacgca aagtaagcct acccacctat 2760
gcttggaagc gtcaacgtta ttggctagag aaacagtccg ctccacgttt agaaacaaca 2820
caagttcgtc ccgcaactgc cattgtagag catcttgaac aaggcaatgt gccgaaaatc 2880
gtggacttgt tagcggcgac ggatgtactt tcaggcgaag cacggaaatt gctaccacagc 2940
atcattgaac tattgggttc aaaacatcgt gaggaagcga cacagaagcc catctgcgat 3000
tggctttatg aagtgggttg gcaaccccag ttgctgacct tatctacctt acctgctgtg 3060
gaaacagagg gttagacaatg gctcatcttc gccgatgcta gtggacacgg tgaagcactt 3120
gcggctcaat tacgtcagca aggggatata attacgcttg tctatgctgg tctaaaatat 3180
cactcggcta ataataaaca aaataccggg ggggacatcc catattttca gattgatccg 3240
atceaaaggg aggattatga aaggttgttt gctgctttgc ctccactgta tggattgttt 3300
catctttgga gtttagatat acttagcttg gacaaagtat ctaacctaat tgaaaatgta 3360
caattaggta gtggcacgct attaaattta atacagacag tcttgcaact tgaaacgccc 3420
acccctagct tgtggctcgt gacaagaac gcgcaagctg tgcgtaaaaa cgatagccta 3480
gtcggagtgc ttcagtcacc cttatggggg atgggtaagg tgatagcctt agaacacctt 3540
gaactcaact gtgtatcaat cgaccttgat ggtgaagggc ttccagatga acaagccaag 3600
tttctggcgg ctgaactccg cgccgcctcc gagttcagac ataccacctt tccccacgaa 3660
agtcaagttg cttggcgtaa taggactcgc tatgtgtcac ggttcaaagg ttatcagaag 3720
catcccgcg cctcatcaaa aatgcctatt cgaccagatg ccacttattt gatcacgggc 3780
ggctttgggt gtttgggctt gcttgtggct cgttggatgg ttgaacaggg ggctacccat 3840
ctatttctga tgggacgcag ccaacccaaa ccagccgccc aaaaacaact gcaagagata 3900
gccgcgctgg gtgcaacagt gacggtgggt caagccgatg ttggcatccg ctcccaagta 3960
gccaatgtgt tggcacagat tgataaggca tatcctttgg ctggtattat tcatactgcc 4020
ggtgtattag acgacggaat cttattgcag caaaattggg cgcgttttag caagggtgtt 4080
gcccccaaac tagagggagc ttggcatcta catacactga ctgaagagat gccgcttgat 4140
ttctttatgt gtttttcttc aacagcagga ttgctgggca gtggtggaca agctaactat 4200
gctgtgcca atgccttttt agatgccttt gcccatcatc ggccaataca aggcttgcca 4260
gctctctcga ttaactggga cgcttgggtc caagtgggaa tgacgggtacg tctccaacaa 4320
gcttcttcac aaagcaccac agttgggcaa gatattagca ctttggaat ttcaccagaa 4380
cagggttgac aaatctttgc ctatcttctg caacaacct ccgccccaat agcggccatt 4440
tctaccgatg ggcttcgcaa gatgtacgac acaagctcgg ccttttttgc tttacttgat 4500
cttgacaggt ctctctccac taccaggag caatctacac tttctcatga agttggcctt 4560
accttactcg aacaattgca gcaagctcgg ccaaaagagc gagagaaaat gttactgcgc 4620
catctacaga cccaagttgc tgcggtcttg cgtagtcccg aactgcccgc agttcatcaa 4680
cccttactcg acttggggat ggattcggtt atgtcacttg aattgatgcg gcgtttgaa 4740
gaaagtctgg ggattcagat gcctgcaacg cttgcattcg attatcctat ggtagaccgt 4800
ttggctaagt ttatactgac tcaaatatgt ataaattctg agccagatac ctcagcagtt 4860
ctcacaccag atggaaatgg ggaggaaaaa gacagtaata aggacagaag taccagcact 4920
tccgttgact caaatattac ttccatggca gaagatttat tcgcactcga atccttacta 4980
aataaaaataa aaagagatca ataa 5004

```

<210> 104

<211> 1667

5 <212> PRT .

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 104

```

Met Ser Gln Pro Asn Tyr Gly Ile Leu Met Lys Asn Ala Leu Asn Glu
1           5           10           15
Ile Asn Ser Leu Arg Ser Gln Leu Ala Val Glu Ala Gln Lys Asn
20          25          30
Glu Ser Ile Ala Ile Val Gly Met Ser Cys Arg Phe Pro Gly Gly Ala
35          40          45
Thr Thr Pro Glu Arg Phe Trp Val Leu Leu Arg Glu Gly Ile Ser Ala
50          55          60

```

10

```

Ile Thr Glu Ile Pro Ala Asp Arg Trp Asp Val Asp Lys Tyr Tyr Asp
65          70          75          80
Ala Asp Pro Thr Ser Ser Gly Lys Met His Thr Arg Tyr Gly Gly Phe
      85          90          95
Leu Asn Glu Val Asp Thr Phe Glu Pro Ser Phe Phe Asn Ile Ala Ala
      100          105          110
Arg Glu Ala Val Ser Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Val
      115          120          125
Ser Trp Glu Ala Leu Glu Ser Gly Asn Ile Val Pro Ala Thr Leu Phe
      130          135          140
Asp Ser Ser Thr Gly Val Phe Ile Gly Ile Gly Gly Ser Asn Tyr Lys
145          150          155          160
Ser Leu Met Ile Glu Asn Arg Ser Arg Ile Gly Lys Thr Asp Leu Tyr
      165          170          175
Glu Leu Ser Gly Thr Asp Val Ser Val Ala Ala Gly Arg Ile Ser Tyr
      180          185          190
Val Leu Gly Leu Met Gly Pro Ser Phe Val Ile Asp Thr Ala Cys Ser
      195          200          205
Ser Ser Leu Val Ser Val His Gln Ala Cys Gln Ser Leu Arg Gln Arg
      210          215          220
Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Gly Leu Leu Ile Asp Pro
225          230          235          240
Asp Glu Met Ile Gly Leu Ser Gln Gly Gly Met Leu Ala Pro Asp Gly
      245          250          255
Ser Cys Lys Thr Phe Asp Ala Asn Ala Asn Gly Tyr Val Arg Gly Glu
      260          265          270
Gly Cys Gly Met Ile Val Leu Lys Arg Leu Ser Asp Ala Thr Ala Asp
      275          280          285
Gly Asp Asn Ile Leu Ala Ile Ile Arg Gly Ser Met Val Asn His Asp
290          295          300
Gly His Ser Ser Gly Leu Thr Ala Pro Arg Gly Pro Ala Gln Val Ser
305          310          315          320
Val Ile Lys Gln Ala Leu Asp Arg Ala Gly Ile Ala Pro Asp Ala Val
      325          330          335
Ser Tyr Leu Glu Ala His Gly Thr Gly Thr Pro Leu Gly Asp Pro Ile
      340          345          350
Glu Met Asp Ser Leu Asn Glu Val Phe Gly Arg Arg Thr Glu Pro Leu
      355          360          365
Trp Val Gly Ser Val Lys Thr Asn Ile Gly His Leu Glu Ala Ala Ser
370          375          380
Gly Ile Ala Gly Leu Ile Lys Val Val Leu Met Leu Lys Asn Lys Gln
385          390          395          400
Ile Pro Pro His Leu His Phe Lys Thr Pro Asn Pro Tyr Ile Asp Trp
      405          410          415
Lys Asn Leu Pro Val Glu Ile Pro Thr Thr Leu His Ala Trp Asp Asp
      420          425          430
Lys Thr Leu Lys Asp Arg Lys Arg Ile Ala Gly Val Ser Ser Phe Ser
      435          440          445
Phe Ser Gly Thr Asn Ala His Ile Val Leu Ser Glu Ala Pro Ser Ser
450          455          460
Glu Leu Ile Ser Asn His Ala Ala Val Glu Arg Pro Trp His Leu Leu
465          470          475          480
Thr Leu Ser Ala Lys Asn Glu Glu Ala Leu Ala Asn Leu Val Gly Leu
      485          490          495
Tyr Gln Ser Phe Ile Ser Thr Thr Asp Ala Ser Leu Ala Asp Ile Cys
500          505          510
Tyr Thr Ala Asn Thr Ala Arg Thr His Phe Ser His Arg Leu Ala Leu

```

[illegible]

Leu	Leu	Pro	Ser	Ile	Ile	Glu	Leu	Leu	Val	Ala	Lys	His	Arg	Glu	Glu
		980					985						990		
Ala	Thr	Gln	Lys	Pro	Ile	Cys	Asp	Trp	Leu	Tyr	Glu	Val	Val	Trp	Gln
		995					1000					1005			
Pro	Gln	Leu	Leu	Thr	Leu	Ser	Thr	Leu	Pro	Ala	Val	Glu	Thr	Glu	
	1010					1015					1020				
Gly	Arg	Gln	Trp	Leu	Ile	Phe	Ala	Asp	Ala	Ser	Gly	His	Gly	Glu	
	1025					1030					1035				
Ala	Leu	Ala	Ala	Gln	Leu	Arg	Gln	Gln	Gly	Asp	Ile	Ile	Thr	Leu	
	1040					1045					1050				
Val	Tyr	Ala	Gly	Leu	Lys	Tyr	His	Ser	Ala	Asn	Asn	Lys	Gln	Asn	
	1055					1060					1065				
Thr	Gly	Gly	Asp	Ile	Pro	Tyr	Phe	Gln	Ile	Asp	Pro	Ile	Gln	Arg	
	1070					1075					1080				
Glu	Asp	Tyr	Glu	Arg	Leu	Phe	Ala	Ala	Leu	Pro	Pro	Leu	Tyr	Gly	
	1085					1090					1095				
Ile	Val	His	Leu	Trp	Ser	Leu	Asp	Ile	Leu	Ser	Leu	Asp	Lys	Val	
	1100					1105					1110				
Ser	Asn	Leu	Ile	Glu	Asn	Val	Gln	Leu	Gly	Ser	Gly	Thr	Leu	Leu	
	1115					1120					1125				
Asn	Leu	Ile	Gln	Thr	Val	Leu	Gln	Leu	Glu	Thr	Pro	Thr	Pro	Ser	
	1130					1135					1140				
Leu	Trp	Leu	Val	Thr	Lys	Asn	Ala	Gln	Ala	Val	Arg	Lys	Asn	Asp	
	1145					1150					1155				
Ser	Leu	Val	Gly	Val	Leu	Gln	Ser	Pro	Leu	Trp	Gly	Met	Gly	Lys	
	1160					1165					1170				
Val	Ile	Ala	Leu	Glu	His	Pro	Glu	Leu	Asn	Cys	Val	Ser	Ile	Asp	
	1175					1180					1185				
Leu	Asp	Gly	Glu	Gly	Leu	Pro	Asp	Glu	Gln	Ala	Lys	Phe	Leu	Ala	
	1190					1195					1200				
Ala	Glu	Leu	Arg	Ala	Ala	Ser	Glu	Phe	Arg	His	Thr	Thr	Ile	Pro	
	1205					1210					1215				
His	Glu	Ser	Gln	Val	Ala	Trp	Arg	Asn	Arg	Thr	Arg	Tyr	Val	Ser	
	1220					1225					1230				
Arg	Phe	Lys	Gly	Tyr	Gln	Lys	His	Pro	Ala	Thr	Ser	Ser	Lys	Met	
	1235					1240					1245				
Pro	Ile	Arg	Pro	Asp	Ala	Thr	Tyr	Leu	Ile	Thr	Gly	Gly	Phe	Gly	
	1250					1255					1260				
Gly	Leu	Gly	Leu	Leu	Val	Ala	Arg	Trp	Met	Val	Glu	Gln	Gly	Ala	
	1265					1270					1275				
Thr	His	Leu	Phe	Leu	Met	Gly	Arg	Ser	Gln	Pro	Lys	Pro	Ala	Ala	
	1280					1285					1290				
Gln	Lys	Gln	Leu	Gln	Glu	Ile	Ala	Ala	Leu	Gly	Ala	Thr	Val	Thr	
	1295					1300					1305				
Val	Val	Gln	Ala	Asp	Val	Gly	Ile	Arg	Ser	Gln	Val	Ala	Asn	Val	
	1310					1315					1320				
Leu	Ala	Gln	Ile	Asp	Lys	Ala	Tyr	Pro	Leu	Ala	Gly	Ile	Ile	His	
	1325					1330					1335				
Thr	Ala	Gly	Val	Leu	Asp	Asp	Gly	Ile	Leu	Leu	Gln	Gln	Asn	Trp	
	1340					1345					1350				
Ala	Arg	Phe	Ser	Lys	Val	Phe	Ala	Pro	Lys	Leu	Glu	Gly	Ala	Trp	
	1355					1360					1365				
His	Leu	His	Thr	Leu	Thr	Glu	Glu	Met	Pro	Leu	Asp	Phe	Phe	Ile	
	1370					1375					1380				
Cys	Phe	Ser	Ser	Thr	Ala	Gly	Leu	Leu	Gly	Ser	Gly	Gly	Gln	Ala	
	1385					1390					1395				
Asn	Tyr	Ala	Ala	Ala	Asn	Ala	Phe	Leu	Asp	Ala	Phe	Ala	His	His	

1400	1405	1410
Arg Arg Ile Gln Gly Leu	Pro Ala Leu Ser Ile	Asn Trp Asp Ala
1415	1420	1425
Trp Ser Gln Val Gly Met	Thr Val Arg Leu Gln	Gln Ala Ser Ser
1430	1435	1440
Gln Ser Thr Thr Val Gly	Gln Asp Ile Ser Thr	Leu Glu Ile Ser
1445	1450	1455
Pro Glu Gln Gly Leu Gln	Ile Phe Ala Tyr Leu	Leu Gln Gln Pro
1460	1465	1470
Ser Ala Gln Ile Ala Ala	Ile Ser Thr Asp Gly	Leu Arg Lys Met
1475	1480	1485
Tyr Asp Thr Ser Ser Ala	Phe Phe Ala Leu Leu	Asp Leu Asp Arg
1490	1495	1500
Ser Ser Ser Thr Thr Gln	Glu Gln Ser Thr Leu	Ser His Glu Val
1505	1510	1515
Gly Leu Thr Leu Leu Glu	Gln Leu Gln Gln Ala	Arg Pro Lys Glu
1520	1525	1530
Arg Glu Lys Met Leu Leu	Arg His Leu Gln Thr	Gln Val Ala Ala
1535	1540	1545
Val Leu Arg Ser Pro Glu	Leu Pro Ala Val His	Gln Pro Phe Thr
1550	1555	1560
Asp Leu Gly Met Asp Ser	Leu Met Ser Leu Glu	Leu Met Arg Arg
1565	1570	1575
Leu Glu Glu Ser Leu Gly	Ile Gln Met Pro Ala	Thr Leu Ala Phe
1580	1585	1590
Asp Tyr Pro Met Val Asp	Arg Leu Ala Lys Phe	Ile Leu Thr Gln
1595	1600	1605
Ile Cys Ile Asn Ser Glu	Pro Asp Thr Ser Ala	Val Leu Thr Pro
1610	1615	1620
Asp Gly Asn Gly Glu Glu	Lys Asp Ser Asn Lys	Asp Arg Ser Thr
1625	1630	1635
Ser Thr Ser Val Asp Ser	Asn Ile Thr Ser Met	Ala Glu Asp Leu
1640	1645	1650
Phe Ala Leu Glu Ser Leu	Leu Asn Lys Ile Lys	Arg Asp Gln
1655	1660	1665

<210> 105

<211> 318

5 <212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 105

ttatgctgca tctaaataga agttccatag ccctgcactg accaacaatca attgatcatc	60
aaaatcggtc acacgattcc tatatgtggg ataaaatttg cagtacagca ggatataaaa	120
tagtttttcc tctatacttc tgagtgtagg cttgcgtccg ccccggggcg cacgtttgcg	180
gtttgctaag gagttgaaca cgggtgcgttc ataggtatca gcaaactgag ataacagctc	240
gttgaatgct tggcgggttaa gtccagtcac tgctcgtagc agtcgctctt gattcaggat	300
gcggtctaag ttcaacat	318

10

<210> 106

<211> 105

<212> PRT

15 <213> *Cylindrospermopsis raciborskii* AWT205

<400> 106

Met Leu Asn Leu Asp Arg Ile Leu Asn Gln Glu Arg Leu Leu Arg Ala	
1 5 10 15	
Met Thr Gly Leu Asn Arg Gln Ala Phe Asn Glu Leu Leu Ser Gln Phe	
20 25 30	

20

Ala Asp Thr Tyr Glu Arg Thr Val Phe Asn Ser Leu Ala Asn Arg Lys
 35 40 45
 Arg Ala Pro Gly Gly Gly Arg Lys Pro Thr Leu Arg Ser Ile Glu Glu
 50 55 60
 Lys Leu Phe Tyr Ile Leu Leu Tyr Cys Lys Phe Tyr Pro Thr Tyr Arg
 65 70 75 80
 Asn Arg Val Thr Asp Phe Asp Asp Gln Leu Met Leu Val Ser Ala Gly
 85 90 95
 Leu Trp Asn Phe Tyr Leu Asp Ala Ala
 100 105

<210> 107

<211> 600

<212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 107

ctactgagtg aaagtgaact tctttccac gtattcgagt agctgttgta agctggcctc 60
 gatggaaagt tccgaagttt ccaccagtaa atctggtgtt ctcggtgggt cgtagggagc 120
 gctaattccc gtaaaagact caatttctcc acggcggtgct tttgcataga gacccttggg 180
 gtcacgttgt tcacaaattt ccacggagtg tgcaatatat acttcatgaa acagatctcc 240
 ggacagaata cggatttgct cccggtcttt cctgtaaggt gaaatgaaag cagtaatcac 300
 taaacaaccc gaatccgcaa aaagtttggc cacctcgcca atacgacgaa tattttccgc 360
 acgatcagca gcagaaaatc ccaagtcagc acataatcca tgacggatat tgtcaccatc 420
 aaggacaaaa gtataccaac ctttctggaa caaatccgc tctaattcta gagccaatgt 480
 tgttttacct gatcctgata atccagtgaa ccatagaatt ccatttcggt gaccattctt 540
 taaacaacga tcaaatgggg acacaagatg tttgtatgt tgaatattgc ttgatttcat 600

<210> 108

<211> 199

<212> PRT

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 108

Met Lys Ser Ser Asn Ile Gln His Thr Lys His Leu Val Ser Pro Phe
 1 5 10 15
 Asp Arg Cys Leu Lys Asn Gly His Arg Asn Gly Ile Leu Trp Phe Thr
 20 25 30
 Gly Leu Ser Gly Ser Gly Lys Thr Thr Leu Ala Leu Glu Leu Glu Arg
 35 40 45
 Ile Leu Phe Gln Lys Gly Trp Tyr Thr Phe Val Leu Asp Gly Asp Asn
 50 55 60
 Ile Arg His Gly Leu Cys Ala Asp Leu Gly Phe Ser Ala Ala Asp Arg
 65 70 75 80
 Ala Glu Asn Ile Arg Arg Ile Gly Glu Val Ala Lys Leu Phe Ala Asp
 85 90 95
 Ser Gly Cys Leu Val Ile Thr Ala Phe Ile Ser Pro Tyr Arg Lys Asp
 100 105 110
 Arg Glu Gln Ile Arg Ile Leu Ser Gly Asp Leu Phe His Glu Val Tyr
 115 120 125
 Ile Ala Thr Pro Met Glu Ile Cys Glu Gln Arg Asp Pro Lys Gly Leu
 130 135 140
 Tyr Ala Lys Ala Arg Arg Gly Glu Ile Glu Ser Phe Thr Gly Ile Ser
 145 150 155 160
 Ala Pro Tyr Glu Pro Pro Arg Thr Pro Asp Leu Leu Val Glu Thr Ser
 165 170 175
 Glu Leu Ser Ile Glu Ala Ser Leu Gln Gln Leu Leu Glu Tyr Val Gly
 180 185 190
 Lys Lys Phe Thr Phe Thr Gln

195

<210> 109

<211> 1548

<212> ADN

<213> *Cylindrospermopsis raciborskii* AWT205

<400> 109

```

atgcctaaat actttaatac tgctggaccc tgtaaattccg aaatccacta tatgctctct 60
cccacagctc gactaccgga tttgaaagca ctaattgacg gagaaaacta ctttataatt 120
cacgcgccgc gacaagtcgg caaaactaca gctatgatag ccttagcacg agaattgact 180
gatagtggaa aatataaccg agttattctt tccgttgaag tgggatcagt attctcccat 240
aatccccagc aagcggagca ggttatttta gaagaatgga aacaggcaat caaattttat 300
ttaccctaaag aactacaacc atcctattgg ccagagcgtg aaacagactc aggaataggc 360
aaaactttta gtgagtggtc cgcacaatct ccaagacctc ttgtaatctt tttacatgaa 420
atcgattccc taacagatga agctttaatc ctaattttta gacaattacg ctcagggttt 480
ccccgtcgtc ctccggggatt tccccattcg gtgggggttaa ttggtatgcg ggatgtgcgg 540
gactataagg ttaaatctgg tggaagtga cgaactgaata cgtcaagtcc tttcaatata 600
aaagcgggaat ccttgacttt aagtaatttc actctgtcag aggtggaaga actttactta 660
caacatacgc aagctacagg acaaattttt accccggaag caattaaaca agcattttat 720
ttaaccgatg ggcaaccatg gttagtaaac gccctagctc gtcaagccac tcagggttta 780
gtgaaagata ttactcaacc cattaccgct gaagtaatta accaagccaa agaagtctcg 840
attcagcgcc aggatacca tttggatagt ttggcagagc gcttacggga agatcgggtc 900
aaagcattat ttcaacctat gttagctgga tcggacttac cagatacccc agaggatgat 960
cgccgtttct tgctagattt aggccttggtt aagcgcagtc ccttgggagg actaaccatt 1020
gccaatccca ttaccagga ggtgattcct cgtgttttgt cccagggtag tcaggatagt 1080
ctaccccaaga ttcaacctac ttggttaaat actgataata ctttaaatcc tgacaaactc 1140
ttaaatgctt tcttagagtt ttggcgacaa catggggaac cattactcaa aagtgcgcct 1200
tatcatgaaa ttgctcccca tttagttttg atggcgtttt tacatcgggt agtgaatgg 1260
ggtggcactt tagaacggga atatgccgtt ggttctggaa gaatggatat ttgtttacgc 1320
tatggcaagg tagtgatggg catagagtta aaggtttggg ggggaaaatc ggatccgtta 1380
acgaagggtt tgaccctaatt ggataaatat ctgggtgggt taggattaga tagaggttgg 1440
ttagtaattt ttgatcaccg tccgggatta ccacccatgg gtgagaggat tagtatggaa 1500
caggccatta gtccagaggg aagaaccatt acagtgattc gtactctag 1548

```

5 <210> 110
 <211> 515
 <212> PRT
 <213> *Cylindrospermopsis raciborskii* AWT205

10 <400> 110

```

Met Pro Lys Tyr Phe Asn Thr Ala Gly Pro Cys Lys Ser Glu Ile His
1      5      10      15
Tyr Met Leu Ser Pro Thr Ala Arg Leu Pro Asp Leu Lys Ala Leu Ile
20      25      30
Asp Gly Glu Asn Tyr Phe Ile Ile His Ala Pro Arg Gln Val Gly Lys
35      40      45
Thr Thr Ala Met Ile Ala Leu Ala Arg Glu Leu Thr Asp Ser Gly Lys
50      55      60
Tyr Thr Ala Val Ile Leu Ser Val Glu Val Gly Ser Val Phe Ser His
65      70      75      80
Asn Pro Gln Gln Ala Glu Gln Val Ile Leu Glu Glu Trp Lys Gln Ala
85      90      95
Ile Lys Phe Tyr Leu Pro Lys Glu Leu Gln Pro Ser Tyr Trp Pro Glu
100     105     110
Arg Glu Thr Asp Ser Gly Ile Gly Lys Thr Leu Ser Glu Trp Ser Ala
115     120     125
Gln Ser Pro Arg Pro Leu Val Ile Phe Leu His Glu Ile Asp Ser Leu
130     135     140

```

Thr Asp Glu Ala Leu Ile Leu Ile Leu Arg Gln Leu Arg Ser Gly Phe
 145 150 155 160
 Pro Arg Arg Pro Arg Gly Phe Pro His Ser Val Gly Leu Ile Gly Met
 165 170 175
 Arg Asp Val Arg Asp Tyr Lys Val Lys Ser Gly Gly Ser Glu Arg Leu
 180 185 190
 Asn Thr Ser Ser Pro Phe Asn Ile Lys Ala Glu Ser Leu Thr Leu Ser
 195 200 205
 Asn Phe Thr Leu Ser Glu Val Glu Glu Leu Tyr Leu Gln His Thr Gln
 210 215 220
 Ala Thr Gly Gln Ile Phe Thr Pro Glu Ala Ile Lys Gln Ala Phe Tyr
 225 230 235 240
 Leu Thr Asp Gly Gln Pro Trp Leu Val Asn Ala Leu Ala Arg Gln Ala
 245 250 255
 Thr Gln Val Leu Val Lys Asp Ile Thr Gln Pro Ile Thr Ala Glu Val
 260 265 270
 Ile Asn Gln Ala Lys Glu Val Leu Ile Gln Arg Gln Asp Thr His Leu
 275 280 285
 Asp Ser Leu Ala Glu Arg Leu Arg Glu Asp Arg Val Lys Ala Ile Ile
 290 295 300
 Gln Pro Met Leu Ala Gly Ser Asp Leu Pro Asp Thr Pro Glu Asp Asp
 305 310 315 320
 Arg Arg Phe Leu Leu Asp Leu Gly Leu Val Lys Arg Ser Pro Leu Gly
 325 330 335
 Gly Leu Thr Ile Ala Asn Pro Ile Tyr Gln Glu Val Ile Pro Arg Val
 340 345 350
 Leu Ser Gln Gly Ser Gln Asp Ser Leu Pro Gln Ile Gln Pro Thr Trp
 355 360 365
 Leu Asn Thr Asp Asn Thr Leu Asn Pro Asp Lys Leu Leu Asn Ala Phe
 370 375 380
 Leu Glu Phe Trp Arg Gln His Gly Glu Pro Leu Leu Lys Ser Ala Pro
 385 390 395 400
 Tyr His Glu Ile Ala Pro His Leu Val Leu Met Ala Phe Leu His Arg
 405 410 415
 Val Val Asn Gly Gly Gly Thr Leu Glu Arg Glu Tyr Ala Val Gly Ser
 420 425 430
 Gly Arg Met Asp Ile Cys Leu Arg Tyr Gly Lys Val Val Met Gly Ile
 435 440 445
 Glu Leu Lys Val Trp Gly Gly Lys Ser Asp Pro Leu Thr Lys Gly Leu
 450 455 460
 Thr Gln Leu Asp Lys Tyr Leu Gly Gly Leu Gly Leu Asp Arg Gly Trp
 465 470 475 480
 Leu Val Ile Phe Asp His Arg Pro Gly Leu Pro Pro Met Gly Glu Arg
 485 490 495
 Ile Ser Met Glu Gln Ala Ile Ser Pro Glu Gly Arg Thr Ile Thr Val
 500 505 510
 Ile Arg Ser
 515

<210> 111

<211> 20

5 <212> ADN

<213> Artificial

<220>

<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* AWT205

10

<400> 111

acttctctcc ttccctatc

20

15 <210> 112

<211> 22

<212> ADN

<213> Artificial

20 <220>

<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* AWT205

<400> 112

	gagtgaaaat gcgtagaact tg	22
5	<210> 113 <211> 22 <212> ADN <213> Artificial	
10	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
	<400> 113	
15	cccaatatct ccctgtaaaa ct	22
20	<210> 114 <211> 20 <212> ADN <213> Artificial	
	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
	<400> 114	
25	tggcaattgt ctctccgtat	20
30	<210> 115 <211> 20 <212> ADN <213> Artificial	
	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
35	<400> 115	
	ctgccgatg aaagtcctct	20
40	<210> 116 <211> 20 <212> ADN <213> Artificial	
45	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
	<400> 116	
50	gcgtgtcgag aaaaaggtgt	20
55	<210> 117 <211> 20 <212> ADN <213> Artificial	
	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
60	<400> 117	
	ctcgacacgc aagaataacg	20
65	<210> 118 <211> 21 <212> ADN	

<213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 5 <400> 118
 atgcttctgc ttggcatgg c 21
 10 <210> 119
 <211> 21
 <212> ADN
 <213> Artificial
 15 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 119
 20 taactcgacg aactttgacc c 21
 <210> 120
 <211> 19
 <212> ADN
 25 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 30 <400> 120
 gccgccaatc ctgcgatg 19
 <210> 121
 35 <211> 22
 <212> ADN
 <213> Artificial
 <220>
 40 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 121
 gaacgtctaa tgttgacag tg 22
 45 <210> 122
 <211> 23
 <212> ADN
 <213> Artificial
 50 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 122
 55 ctggtacgta gtcgcaaagg tgg 23
 <210> 123
 <211> 26
 60 <212> ADN
 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 65 <400> 123

ctgacggtac atgtatttc tgtgac 26

5 <210> 124
<211> 30
<212> ADN
<213> Artificial

10 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 124

15 cgtctcatat gcagatctta ggaatttcag 30

20 <210> 125
<211> 25
<212> ADN
<213> Artificial

25 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 125

30 gcttactacc acgatagtc tgccg 25

35 <210> 126
<211> 22
<212> ADN
<213> Artificial

40 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 126

45 tctatgttta gcaggtggtg tc 22

50 <210> 127
<211> 20
<212> ADN
<213> Artificial

55 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 127

60 ttctgaaga cgagccataa 20

65 <210> 128
<211> 20
<212> ADN
<213> Artificial

70 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 128

75 ggctcgccgc ggacattaaa 20

80 <210> 129
<211> 20
<212> ADN

<213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 5 <400> 129
 atgctaagtc ggtgggagta 20
 10 <210> 130
 <211> 20
 <212> ADN
 <213> Artificial
 15 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 130
 20 aaagcagttc cgacgacatt 20
 <210> 131
 <211> 23
 <212> ADN
 25 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 30 <400> 131
 cctatttcga ttattgtttt cgg 23
 <210> 132
 35 <211> 20
 <212> ADN
 <213> Artificial
 <220>
 40 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 132
 gataccgatc ataaactacg 20
 45 <210> 133
 <211> 21
 <212> ADN
 <213> Artificial
 50 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3 sequence
 <400> 133
 55 gcaaattttg caggagtaat g 21
 <210> 134
 <211> 21
 60 <212> ADN
 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 65 <400> 134

gcaaatttg caggagtaat g 21

5 <210> 135
<211> 23
<212> ADN
<213> Artificial

10 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 135

15 tttgggtaa actttatagc cat 23

20 <210> 136
<211> 22
<212> ADN
<213> Artificial

25 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 136

30 tgggtctgga cagttgtaga ta 22

35 <210> 137
<211> 23
<212> ADN
<213> Artificial

40 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 137

45 aaggggaaaa caaaattatc aat 23

50 <210> 138
<211> 20
<212> ADN
<213> Artificial

55 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 138

60 ggcatcgcc tgctaaaaat 20

65 <210> 139
<211> 23
<212> ADN
<213> Artificial

<220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 139

cctcatttc atttctagac gtt 23

65 <210> 140
<211> 20
<212> ADN

<213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 5 <400> 140
 ccacttcaac taaaacagca 20
 10 <210> 141
 <211> 20
 <212> ADN
 <213> Artificial
 15 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 141
 20 aaaaattttg gaggggtagc 20
 <210> 142
 <211> 20
 <212> ADN
 25 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 30 <400> 142
 atccaagatg cgacaacact 20
 <210> 143
 35 <211> 21
 <212> ADN
 <213> Artificial
 <220>
 40 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 143
 ggtccttgcg cagatagagt g 21
 45 <210> 144
 <211> 21
 <212> ADN
 <213> Artificial
 50 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 144
 55 cactctatct gcgcaaggac c 21
 <210> 145
 <211> 21
 60 <212> ADN
 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 65 <400> 145

	tgactgcatt cgctgtataa a	21
5	<210> 146 <211> 22 <212> ADN <213> Artificial	
10	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
	<400> 146	
15	ttcataagac ggctgttgaa tc	22
20	<210> 147 <211> 30 <212> ADN <213> Artificial	
	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
	<400> 147	
25	ctcgagttaa aaaagagtgat aaatgaaagg	30
30	<210> 148 <211> 23 <212> ADN <213> Artificial	
35	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
	<400> 148	
40	ttctataact gctgccaat ttt	23
45	<210> 149 <211> 23 <212> ADN <213> Artificial	
	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
	<400> 149	
50	aattttggag tgactggtaa tgg	23
55	<210> 150 <211> 23 <212> ADN <213> Artificial	
	<220> <223> Basado en la secuencia de <i>Cylindrospermopsis raciborskii</i> T3	
60	<400> 150	
	ccataaccag tcaactccaaa att	23
65	<210> 151 <211> 21 <212> ADN	

<213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 5 <400> 151
 ttttagttgt tacttttggc g 21
 10 <210> 152
 <211> 20
 <212> ADN
 <213> Artificial
 15 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 152
 20 acagcagatg agagaaagta 20
 <210> 153
 <211> 20
 <212> ADN
 25 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 30 <400> 153
 gggttgtctt gctgatttc 20
 <210> 154
 35 <211> 22
 <212> ADN
 <213> Artificial
 <220>
 40 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 154
 cattaaaata agtccggaca gg 22
 45 <210> 155
 <211> 20
 <212> ADN
 <213> Artificial
 50 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 155
 55 ttaaacagaa tgaggagcaa 20
 <210> 156
 <211> 20
 60 <212> ADN
 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 65 <400> 156

aaacaacaca cccatctaag 20

5 <210> 157
<211> 20
<212> ADN
<213> Artificial

10 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 157

15 ttaataaggc atccccaaga 20

20 <210> 158
<211> 20
<212> ADN
<213> Artificial

25 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 158

30 gaaatggctg tgtaaaaact 20

35 <210> 159
<211> 20
<212> ADN
<213> Artificial

40 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 159

45 tctgccatat ccccaacct 20

50 <210> 160
<211> 20
<212> ADN
<213> Artificial

55 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 160

60 gatcgcccg caggaagact 20

65 <210> 161
<211> 20
<212> ADN
<213> Artificial

<220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3

60 <400> 161

tccggcttga cctgtggac 20

65 <210> 162
<211> 20
<212> ADN

<213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 5 <400> 162
 tgcgatgatt ttgcctctgt 20
 10 <210> 163
 <211> 20
 <212> ADN
 <213> Artificial
 15 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 163
 20 aaaatttgca caccacacg 20
 <210> 164
 <211> 27
 <212> ADN
 25 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 30 <400> 164
 ttggattgaa cgtgtaattg aaaaagc 27
 <210> 165
 35 <211> 27
 <212> ADN
 <213> Artificial
 <220>
 40 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 165
 gcttttcaa ttacacgttc aatccaa 27
 45 <210> 166
 <211> 19
 <212> ADN
 <213> Artificial
 50 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 166
 55 aaatggcgta tcgactaac 19
 <210> 167
 <211> 21
 60 <212> ADN
 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 65 <400> 167

atataggagc gcataaagtg c 21

5 <210> 168
<211> 20
<212> ADN
<213> Artificial

10 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 168

15 cttggataa gtctgtgat 20

20 <210> 169
<211> 20
<212> ADN
<213> Artificial

25 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 169

30 aacactcatt agattcatct 20

35 <210> 170
<211> 21
<212> ADN
<213> Artificial

40 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 170

45 tccactaaat cctttgaatt g 21

50 <210> 171
<211> 21
<212> ADN
<213> Artificial

55 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 171

60 tgttgtctg gatcgatcc t 21

65 <210> 172
<211> 20
<212> ADN
<213> Artificial

gcagttcagg tccatgaaac 20

65 <210> 173
<211> 20
<212> ADN

<213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 5 <400> 173
 agcccagtca caaccttcgt 20
 10 <210> 174
 <211> 21
 <212> ADN
 <213> Artificial
 15 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 174
 20 tctggaagta ctgcactgt c 21
 <210> 175
 <211> 22
 <212> ADN
 25 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 30 <400> 175
 tgtaactccg tcaggacata aa 22
 <210> 176
 35 <211> 23
 <212> ADN
 <213> Artificial
 <220>
 40 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 176
 tgcaaatttt agtagcaata acg 23
 45 <210> 177
 <211> 27
 <212> ADN
 <213> Artificial
 50 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 177
 55 ctttactaat tatagcgggg atattat 27
 <210> 178
 <211> 20
 60 <212> ADN
 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 65 <400> 178

cagtggggaa atagatggat 20

5 <210> 179
<211> 20
<212> ADN
<213> Artificial

10 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 179

15 tggtcataaa agcgggattc 20

20 <210> 180
<211> 18
<212> ADN
<213> Artificial

25 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 180

30 ggatcttggc gcaattta 18

35 <210> 181
<211> 23
<212> ADN
<213> Artificial

40 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 181

45 gtagagact tggaacgtat tgg 23

50 <210> 182
<211> 19
<212> ADN
<213> Artificial

55 <220>
<223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
<400> 182

60 ccaaaccag aagaaatcc 19

65 <210> 183
<211> 22
<212> ADN
<213> Artificial

aatctatagc caaaaccct aa 22

65 <210> 184
<211> 19
<212> ADN

<213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 5 <400> 184
 actgtgtgaa caattcccc 19
 10 <210> 185
 <211> 29
 <212> ADN
 <213> Artificial
 15 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 <400> 185
 20 gcaacaagac tacatttagt agatttaga 29
 <210> 186
 <211> 27
 <212> ADN
 25 <213> Artificial
 <220>
 <223> Basado en la secuencia de *Cylindrospermopsis raciborskii* T3
 30 <400> 186
 gcttttcaa ttacacgttc aatccaa 27

REIVINDICACIONES

1. Un método para detectar un organismo cianotóxico que comprende las etapas de obtener una muestra para usarse en el método y analizar la muestra para la presencia de una agrupación de genes SXT presente sólo en organismos productores de saxitoxina, en el que dicho análisis comprende detectar:

5 (i) un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, y SEQ ID NO: 36;

(ii) un polinucleótido variante que tiene al menos 80% de identidad de secuencia con un polinucleótido de (i);

(iii) un ácido ribonucleico o ADN complementario codificado por una secuencia según (i);

10 (iv) un polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, y SEQ ID NO: 37; o,

(v) un polipéptido variante que tiene al menos 80% de identidad de secuencia con un polipéptido de (iv),

y en el que dicha presencia es indicativa de organismos cianotóxicos en la muestra.

2. El método según la reivindicación 1, en el que dicho organismo cianotóxico es una cianobacteria o un dinoflagelado.

15 3. El método según la reivindicación 1, en el que dicho análisis comprende la amplificación de ADN de la muestra por la reacción en cadena de la polimerasa utilizando uno o más cebadores que comprenden una secuencia seleccionada del grupo que consiste en SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de cebador.

25 4. El método según una cualquiera de las reivindicaciones 1 a 3, que comprende además analizar la muestra para la presencia de uno o más de:

30 (i) un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polinucleótido,

(ii) un ácido ribonucleico o ADN complementario codificado por una secuencia según (i),

35 (iii) un polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, y SEQ ID NO: 110, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polipéptido.

40 5. El método según la reivindicación 4, en el que dicho análisis comprende la amplificación de ADN de la muestra por la reacción en cadena de la polimerasa utilizando uno o más cebadores que comprenden una secuencia seleccionada del grupo que consiste en SEQ ID NO: 111, SEQ ID NO: 112, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias.

6. Un kit para la detección de organismos cianotóxicos, comprendiendo el kit al menos un agente para detectar la presencia de una agrupación de genes SXT presente sólo en organismos productores de saxitoxina, en el que bien:

45 (i) dicho agente puede detectar un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polinucleótido y dicho agente es un cebador o una sonda;

(ii) dicho agente puede detectar un ácido ribonucleico o ADN complementario codificado por una secuencia según (i) y dicho agente es un cebador o una sonda; o

50 (iii) dicho agente puede detectar un polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 37, y variantes que tienen al

menos 80% de identidad de secuencia con una cualquiera de dichas secuencia de polipéptido, y dicho agente es un anticuerpo capaz de unirse específicamente al polipéptido.

7. El kit según la reivindicación 6, en el que dicho al menos un agente es un cebador o sonda que comprende una secuencia de nucleótidos seleccionada del grupo que consiste en SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de nucleótidos.

8. El kit según la reivindicación 6 o reivindicación 7, que comprende además al menos un agente adicional para detectar la presencia de uno o más de:

(i) un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polinucleótido,

(ii) un ácido ribonucleico o ADN complementario codificado por una secuencia según (i), o,

(iii) un polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92, SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, y SEQ ID NO: 110, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polipéptido.

9. El kit según la reivindicación 8, en el que dicho al menos un agente adicional es un anticuerpo, cebador o sonda, en el que dicho cebador o sonda comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO: 111, SEQ ID NO: 112, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias.

10. Un polinucleótido aislado que comprende una secuencia de nucleótidos que comparte al menos 80% de identidad de secuencia con SEQ ID NO: 1 o una secuencia de nucleótidos que comparte al menos 80% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, y SEQ ID NO: 68.

11. Un ácido ribonucleico aislado o un ADN complementario aislado codificado por una secuencia según la reivindicación 10.

12. Un polipéptido aislado de la ruta biosintética de la saxitoxina codificado por un fragmento según la reivindicación 10 y que comprende una secuencia de aminoácidos que comparte al menos 80% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, y SEQ ID NO: 69.

13. Una sonda o cebador que hibrida específicamente con uno o más de:

(i) un polinucleótido según la reivindicación 10, o

(ii) un ácido ribonucleico o ADN complementario según la reivindicación 11,

en el que dicho cebador o sonda comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, y variantes que tienen al menos 80% de identidad de secuencia con una cualquiera de dichas secuencias de polinucleótido.

14. Un anticuerpo aislado capaz de unirse específicamente a un polipéptido según la reivindicación 12.

Cepas de Cianobacteria	Toxicidad (Ref)	<i>sxtA</i>	<i>sxtG</i>	<i>sxtH</i>	<i>sxtI</i>	<i>sxtX</i>
<i>A. circinalis</i> AWQC118C	PSP (54)	+	+	+	+	-
<i>A. circinalis</i> AWQC131C	PSP (25)	+	+	+	+	-
<i>A. circinalis</i> AWQC134C	PSP (54)	+	+	+	+	-
<i>A. circinalis</i> AWQC150E	PSP (54)	+	+	+	+	-
<i>A. circinalis</i> AWQC173A	PSP (54)	+	+	+	+	-
<i>A. circinalis</i> AWQC271C	- (54)	-	-	-	-	-
<i>A. circinalis</i> AWQC306A	- (54)	-	-	-	-	-
<i>A. circinalis</i> AWQC310F	- (54)	-	-	-	-	-
<i>A. circinalis</i> AWQC342D	- (54)	-	-	-	-	-
<i>Aph. flos-aquae</i> NH-5	PSP (26)	+	+	+	+	+
<i>Aph. ovalisporum</i> APH028A	CYLN (46)	-	-	-	-	-
<i>C. raciborskii</i> T3	PSP (23)	+	+	+	+	+
<i>C. raciborskii</i> 23B	CYLN (58)	-	-	-	-	-
<i>C. raciborskii</i> GOON	CYLN (43)	-	-	-	-	-
<i>C. raciborskii</i> GERM1	- (30)	-	-	-	-	-
<i>C. raciborskii</i> MARAU1	- (30)	-	-	-	-	-
<i>L. wollei</i>	PSP (7)	+	+	+	+	+

FIGURA 1A

Cebador	Desde	Hasta	Dirección	Secuencia	Gen	
SEQ ID NO: 133	1917	1937	→	GCATAATTTGACAGGAGTAATG	esterol desaturasa	sxtD
SEQ ID NO: 134	2744	2763	→	AGAGAIGCTATGCTTTTCAA	InsB	orf3
SEQ ID NO: 135	2889	2911	→	TTTGGGTAACCTTTATAGCCAT	InsB	orf3
SEQ ID NO: 136	3020	3041	←	TGGGTCGGACAGTTGTAGATA	InsA	orf4
SEQ ID NO: 137	3306	3328	←	AAGGGAAACAAATAATTATCAAT	InsA	orf4
SEQ ID NO: 138	3396	3415	→	GCGGATCGCCCTGCTAAAAAT	InsA	orf4
SEQ ID NO: 139	3717	3739	→	CCCTATTTCATTTCTAGACGTT	SPUR	sxtC
SEQ ID NO: 140	4201	4220	→	CCACTTCAACTAAACACAGCA	citidina desaminasa	sxtB
SEQ ID NO: 141	4362	4381	←	AAAAAATTTGGAGGGGTAGC	citidina desaminasa	sxtB
SEQ ID NO: 142	4932	4951	→	ATCCAAGATGCGACAAACACT	citidina desaminasa	sxtB
SEQ ID NO: 70	5193	5212	→	TTAATTGCTTGGTCTAATCTC	PKS	sxtA
SEQ ID NO: 71	5206	5225	←	CAATACCGAAGAGGAGATAG	PKS	sxtA
SEQ ID NO: 72	5345	5364	←	TAGGCGTGTAGTGGGAGAT	PKS	sxtA
SEQ ID NO: 73	5415	5434	→	TGTGTAACCAATTGTGTAGT	PKS	sxtA
SEQ ID NO: 74	5478	5497	←	TTAGCCGGATTACAGGTGAA	PKS	sxtA
SEQ ID NO: 75	6136	6155	←	CTGGACTCGGCTTGTGCTT	PKS	sxtA
SEQ ID NO: 76	6933	6952	→	CAGCGAGTTACACCCACCAC	PKS	sxtA
SEQ ID NO: 76	7035	7054	←	CTCGCACTAAATATCTACC	PKS	sxtA
SEQ ID NO: 78	7434	7452	→	AAAACCTCAGCTTCCACAA	PKS	sxtA
SEQ ID NO: 79	7537	7558	←	ATGATTTGGAGGTCCATIGIT	PKS	sxtA
SEQ ID NO: 113	7820	7841	→	CCCATAATCTCCCTGTAAAACI	PKS	sxtA
SEQ ID NO: 114	8170	8189	←	IGGCAATTGTCTCTCCGTAI	PKS	sxtA
SEQ ID NO: 115	8742	8761	←	CTGCCGATGAAGAATCCICT	PKS	sxtA
SEQ ID NO: 116	8772	8791	←	GCGTGTGAGAAAGAAGTGT	PKS	sxtA
SEQ ID NO: 117	8782	8801	→	CTCGACACCAAGAATAACG	PKS	sxtA
SEQ ID NO: 143	9390	9410	→	GGTCTTGGCAGATAGAGTG	semejante a chaperona	sxtE
SEQ ID NO: 144	9390	9410	←	CACCTATCTGCGCAAGGACC	semejante a chaperona	sxtE
SEQ ID NO: 145	9856	9876	←	TGACTCGATTCGCTGTATAAA	MATE I	sxtF
SEQ ID NO: 118	10080	10100	→	ATGCTTCTGCTTTGGCATGGC	amidinotransferasa	sxtG
SEQ ID NO: 119	11468	11488	←	TAACTCGACGAACATTGACCC	amidinotransferasa	sxtG

FIGURA 1B

Cebador	Desde	Hasta	Dirección	Secuencia	Gen	
SEQ ID NO: 120	11551	11569	→	GCCGCCAATCCTCGCGATG	amidontransferasa	sxtG
SEQ ID NO: 121	12256	12277	←	GAACGTCTAATGTTGCACAGTG	amidontransferasa	sxtG
SEQ ID NO: 122	12410	12432	→	CTGGTACGTAGTCGCAAGGTGG	dioxigenasa I	sxtH
SEQ ID NO: 123	13292	13317	→	CTGACGGTACATGTATTTCTCTGTGAC	dioxigenasa I	sxtH
SEQ ID NO: 124	13540	13561	→	cgctcatATGCAGATCTTAGGAATTTTCAG	carbamoltransferasa	sxtL
SEQ ID NO: 125	13561	13585	→	GCCTTACTACCCAGCATAGTGTGCGG	carbamoltransferasa	sxtL
SEQ ID NO: 126	14451	14472	→	TCTATGTTTAGCAGGTGGTGTGTC	carbamoltransferasa	sxtL
SEQ ID NO: 127	14735	14754	←	TTCTGCAAGACGAGCCATAA	carbamoltransferasa	sxtL
SEQ ID NO: 128	15211	15230	←	GGTTGCGCGCGGACATTAA	carbamoltransferasa	sxtL
SEQ ID NO: 146	15709	15730	→	TTCCATAAGACGGCTGTTGAATC	proteína hipotética	sxtJ
SEQ ID NO: 147	15966	15989	←	ctcgagTTAAAAAAGAGTGTAAATGAAAGG	proteína hipotética	sxtK
SEQ ID NO: 148	16326	16348	←	TTCTATAACTGCTGCCAAATTTT	GDSL-lipasa	sxtL
SEQ ID NO: 149	16400	16422	→	AATTTTGGAGTGAAGTGGTTATGG	GDSL-lipasa	sxtL
SEQ ID NO: 150	16400	16422	←	CCATAACCACTCACTCCAAAATT	GDSL-lipasa	sxtL
SEQ ID NO: 151	16929	16949	→	TTTTAGTTGTTACTTTTGGCG	GDSL-lipasa	sxtL
SEQ ID NO: 152	17215	17234	→	ACAGCAGATGAGAGAAAAGTA	GDSL-lipasa	sxtL
SEQ ID NO: 153	18054	18073	→	GGGTGTCTTGTGATTTTC	MATE II	sxtM
SEQ ID NO: 154	18721	18742	←	CATTAAAATAAGTCCGGACAGG	MATE II	sxtM
SEQ ID NO: 155	19133	19152	←	TTAAACAGAAATGAGGAGCAA	MATE II	sxtM
SEQ ID NO: 156	19260	19279	←	AAACAACACACCCATCTAAG	sulfotransferasa	sxtN
SEQ ID NO: 157	19531	19550	→	TTAATAAGGCATCCCCAAGA	sulfotransferasa	sxtN
SEQ ID NO: 158	19728	19747	←	GAAATGGCTGTGTAAAAAAT	sulfotransferasa	sxtN
SEQ ID NO: 129	20584	20603	←	ATGCTAATGCGGTGGGAGTA	cefalosporin hidroxilasa	sxtX
SEQ ID NO: 130	20643	20662	→	AAAGCAGTTCCGACGACATT	cefalosporin hidroxilasa	sxtX
SEQ ID NO: 131	20831	20853	→	CCTATTTTCGATTATTGTTTTCGG	cefalosporin hidroxilasa	sxtX
SEQ ID NO: 132	21252	21271	←	GATACCGATCATAAACTACG	cefalosporin hidroxilasa	sxtX
SEQ ID NO: 159	21290	21309	→	TCTGCCATATCCCCAACCTA	ferredoxina	sxtW
SEQ ID NO: 160	21445	21464	←	GATCGCCCCGACAGGAAGACT	ferredoxina	sxtW
SEQ ID NO: 161	22020	22039	→	TCCGGCTTGACCTGCTGGAC	succinato deshidrogenasa	sxtV

FIGURA 1B (cont)

Cebador	Desde	Hasta	Dirección	Secuencia	Gen	
SEQ ID NO: 162	22715	22734	-->	TGCGATGATTTTGCTCTGT	succinato deshidrogenasa	sxtV
SEQ ID NO: 163	22801	22820	-->	AAAATTTGCACCCACACG	succinato deshidrogenasa	sxtV
SEQ ID NO: 164	22942	22968	-->	TTGGATTGAACGTGTAAATTGAAAAAGC	succinato deshidrogenasa	sxtV
SEQ ID NO: 186	22942	22968	<--	GCITTTTCAATTACACGTTCAATCCAA	succinato deshidrogenasa	sxtV
SEQ ID NO: 165	23434	23453	<--	GTTTAGTCGATACGCCATT	succinato deshidrogenasa	sxtV
SEQ ID NO: 166	23434	23453	-->	AAATGGCGTATCGACTAAC	succinato deshidrogenasa	sxtV
SEQ ID NO: 167	24095	24115	-->	ATATAGGAGCGCATAAAGTGC	succinato deshidrogenasa	sxtV
SEQ ID NO: 168	24728	24747	-->	CTTGGTATAAGTCTTGAT	dioxigenasa II	sxtT
SEQ ID NO: 169	25426	25445	<--	AACACTCATTAGATTCATCT	fitanoil-CoA dioxigenasa	sxtS
SEQ ID NO: 170	25979	25999	<--	TCCACTAAATCCTTTGAATTG	fitanoil-CoA dioxigenasa	sxtS
SEQ ID NO: 171	26279	26299	-->	TGTTTGTCTGGATCGGATCCT	proteína desconocida	orf24
SEQ ID NO: 172	26451	26470	-->	GCAGTTCAGTCCATGAAAC	proteína desconocida	orf25
SEQ ID NO: 173	27155	27174	-->	AGCCCAAGTCACAACCTTCGT	GNAT transferase	sxtR
SEQ ID NO: 174	27508	27528	-->	TCTGGAAGTACTTGCACTGTC	proteína desconocida	sxtQ
SEQ ID NO: 175	28197	28218	-->	TGTAACTCCGTCAGGACATAAA	proteína desconocida	sxtQ
SEQ ID NO: 176	28395	28417	<--	TGCAAAATTTTAGTAGCAATAACG	semejante a toxina RTX	sxtP
SEQ ID NO: 177	29532	29558	<--	CTTTACTAATTATAGCGGGATATTAT	semejante a toxina RTX	sxtP
SEQ ID NO: 178	29868	29887	<--	CAGTGGGGAATAGATGGAT	adenilil sulfato quinasa	sxtO
SEQ ID NO: 179	30249	30268	<--	TGGTCATAAAAGCGGGATTC	adenilil sulfato quinasa	sxtO
SEQ ID NO: 180	31745	31762	-->	GGATCTTGGCGCAATTA	IS4	orf29
SEQ ID NO: 181	33031	33053	<--	GTTAGAGACTTGGAACTGATTGG	PhoU	sxtY
SEQ ID NO: 182	34711	34729	-->	CCAAACCCAGAGAAATCC	histidina quinasa	sxtZ
SEQ ID NO: 183	35100	35121	-->	AATCTATAGCCAAACCCCTAA	ribótido isomerasa	
SEQ ID NO: 184	36447	36465	-->	ACTGTGTGAACAATTCCCC	ribótido isomerasa	
SEQ ID NO: 185	36652	36680	-->	GCAACAAGACTACATTAGTAGATTAGA	ribótido isomerasa	

FIGURA 1B (cont)

Nombre	Familia de Enzimas	Tamaño (pé)	Concordancia de Similitud Blast	(%)	Función Posible
<i>orf1</i>	proteína desconocida	1320	BAB76734.1 <i>Nazot</i> PCC7120	32	desconocida
<i>orf2</i>	semejante a esteroil desaturasa	759	ABG52264.1 <i>Trichodesmium erythraeum</i>	63	desaturación
<i>orf3</i>	transposasa InsB	392	CAE11915.2 <i>Microcystis aeruginosa</i>	36	transposición
<i>orf4</i>	transposasa InsA	360	CAE11914.1 <i>Microcystis aeruginosa</i>	71	transposición
<i>orf5</i>	proteína desconocida	354	no similarity found		reguladora
<i>orf6</i>	citidina desaminasa	957	EAS64681.1 <i>Vibrio angustum</i>	62	ciclación
<i>orf7</i>	metiltransferasa	1506	ABF89568.1 <i>Myxococcus xanthus</i>	64	metilación
	GNAT	633	AAT70096.1 <i>CurtA Lyngbya majuscula</i>	64	carga de ACP
	proteína transportadora de acilo	324	AAV97870 <i>Omb Theonella swinhoei</i>	59	ACP
	AONS	1275	ABD13093.1 <i>Frankia</i> sp. Cc13	61	condensación de Chaisen
<i>orf8</i>	proteína desconocida	387	ABE53436.1 <i>Shewanella denitrificans</i>	52	desconocida
<i>orf9</i>	MATE	1416	NatM ABC44739.1 <i>Salinibacter ruber</i>	52	exportación de PST
<i>orf10</i>	amidinotransferasa	1134	ABA05575.1 <i>Nitrobacter vinogradskyi</i>	71	amidinotransferencia
<i>orf11</i>	fenilpropanato dioxygenasa	1005	ZP_00243439.1 <i>Rubrivivax gelatinosus</i>	50	hidroxilación C-12
<i>orf12</i>	carbamoltransferasa	1839	ABG50968.1 <i>Trichodesmium erythraeum</i>	32	carbamolación

FIGURA 2

<i>xtxJ</i>	proteína desconocida	444	EAM51043.1 <i>Crocospiraera wazzenii</i>	72	reguladora
<i>xtxK</i>	proteína desconocida	165	ABG50954.1 <i>Trichodesmium erythraeum</i>	81	reguladora
<i>xtxL</i>	GDSL-lipasa	1299	ABG50952.1 <i>Trichodesmium erythraeum</i>	60	ciclación
<i>xtxM</i>	MATE	1449	NorM ABC44739.1 <i>Salinibacter ruber</i>	53	exportación de PST
<i>xtxN</i>	sulfotransferasa	831	ABG53102.1 <i>Trichodesmium erythraeum</i>	57	sulfotransferencia
<i>xtxX</i>	cefalosporin hidroxilasa	774	ABG50679.1 <i>Trichodesmium erythraeum</i>	77	II-hidroxilación
<i>xtxW</i>	ferredoxina	327	ZP_00106179.2 <i>Nostoc punctiforme</i>	99	portador de electrones
<i>xtxY</i>	succinato deshidrogenasa	1653	ABA24604.1 <i>Anabaena variabilis</i>	92	dioxigenasa reductasa
<i>xtxU</i>	alcohol deshidrogenasa	750	ZP_00111652.1 <i>Nostoc punctiforme</i>	83	reducción de C-1
<i>xtxT</i>	fenilpropionato dioxigenasa	1005	ZP_00243439.1 <i>Rubrivivax gelatinosus</i>	48	hidroxilación de C-12
<i>xtxS</i>	flavonil-CoA dioxigenasa	726	ABG30370.1 <i>Roseobacter denitrificans</i>	41	formación de anillo
<i>orf24</i>	proteína desconocida	576	no similarity found		desconocida
<i>xtxR</i>	acil transferasa	777	AAU76161.1 <i>Legionella pneumophila</i>	54	desconocida
<i>xtxQ</i>	proteína desconocida	777	EARG4935.1 <i>Bacillus</i> sp. NRRL B-14911	46	desconocida
<i>xtxP</i>	toxina RTX	1227	ABA20206.1 <i>Anabaena variabilis</i>	68	unión de PST
<i>xtxO</i>	adenil sulfato quinasa	603	ZP_00053494.2 <i>Magnetospirillum magnetotacticum</i>	76	biosíntesis de PAPS
<i>orf29</i>	transposasa, IS4	1350	EA022567.1 <i>Synprokocobacter fumaroxidans</i>	61	transposición
<i>xtxY</i>	PboU	666	BAB76200.1 <i>Nostoc PCC7120</i>	87	transducción de la señal
<i>xtxZ</i>	histidina quinasa	1353	ABA22975.1 <i>Anabaena variabilis</i>	78	transducción de la señal
<i>ompR</i>	OmpR	819	ZP_00108178.2 <i>Nostoc punctiforme</i>	91	transducción de la señal
<i>hixA</i>	PROFAR isomerasa	774	ABA22979.1 <i>Anabaena variabilis</i>	90	biosíntesis de histidina
<i>orf34</i>	proteína desconocida	396	ZP_00345366.1 <i>Nostoc punctiforme</i>	84	desconocida

FIGURA 2 (cont)

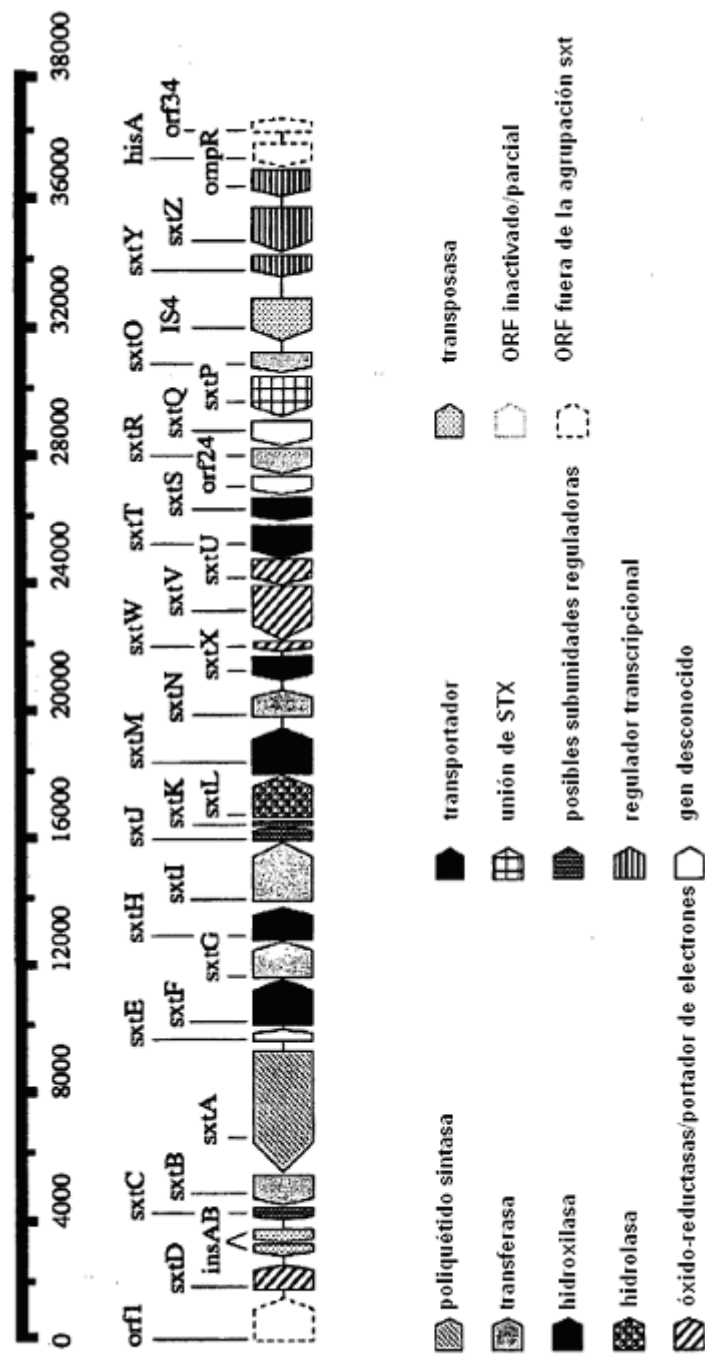
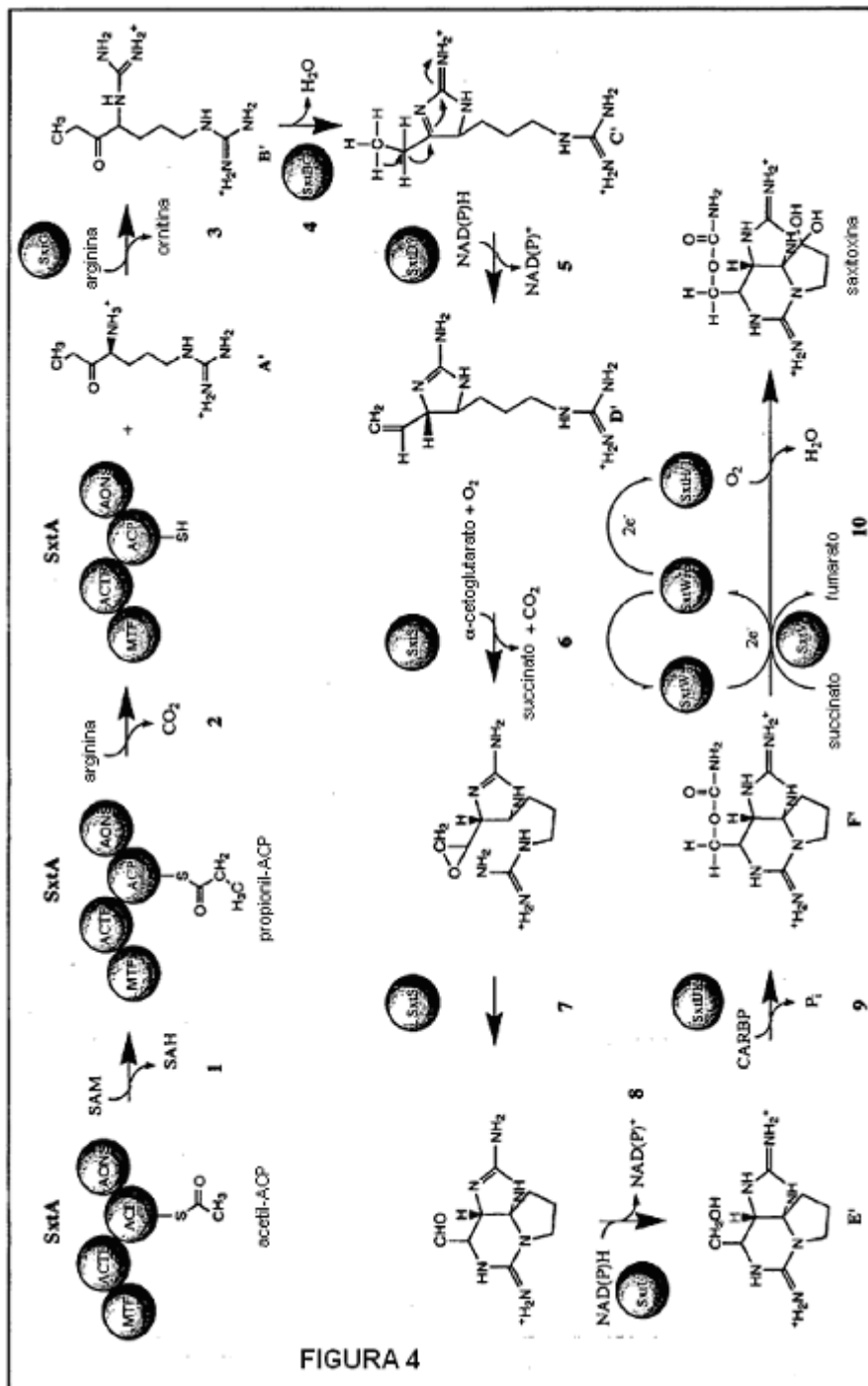


FIGURA 3



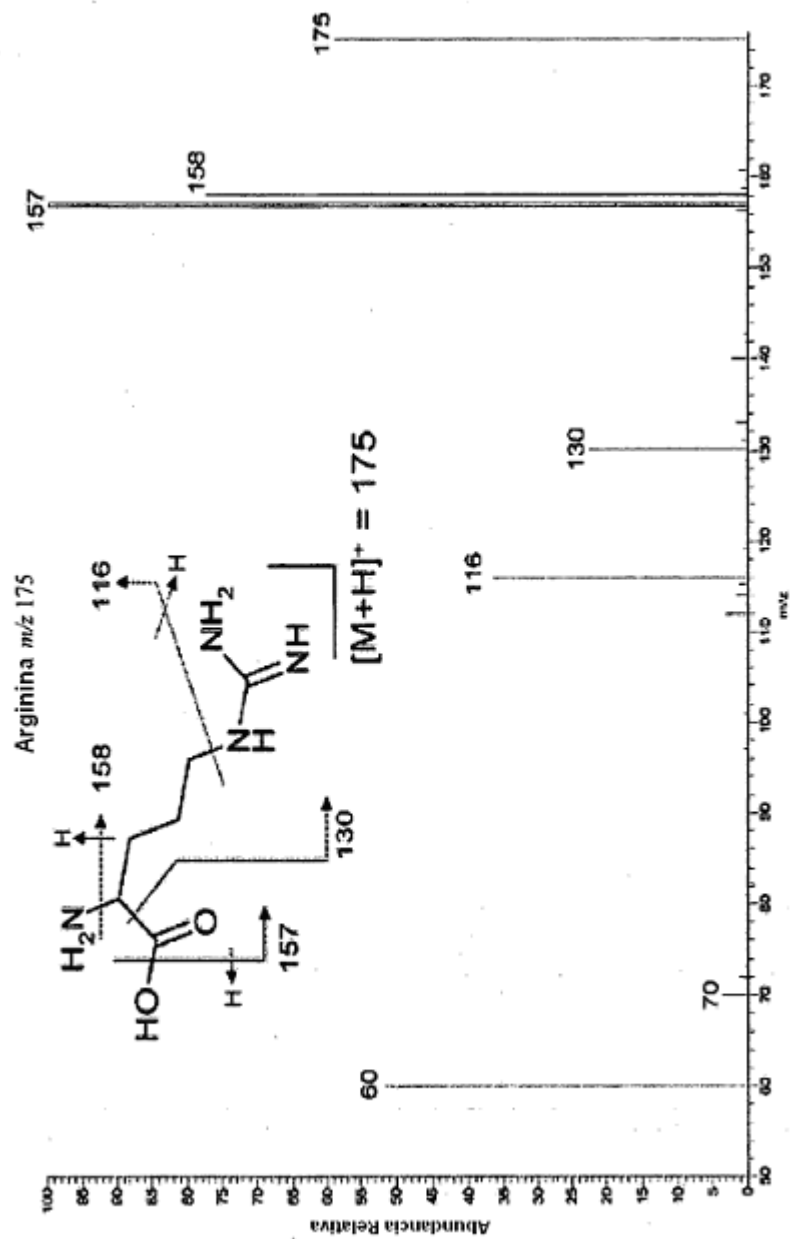


FIGURA 5A

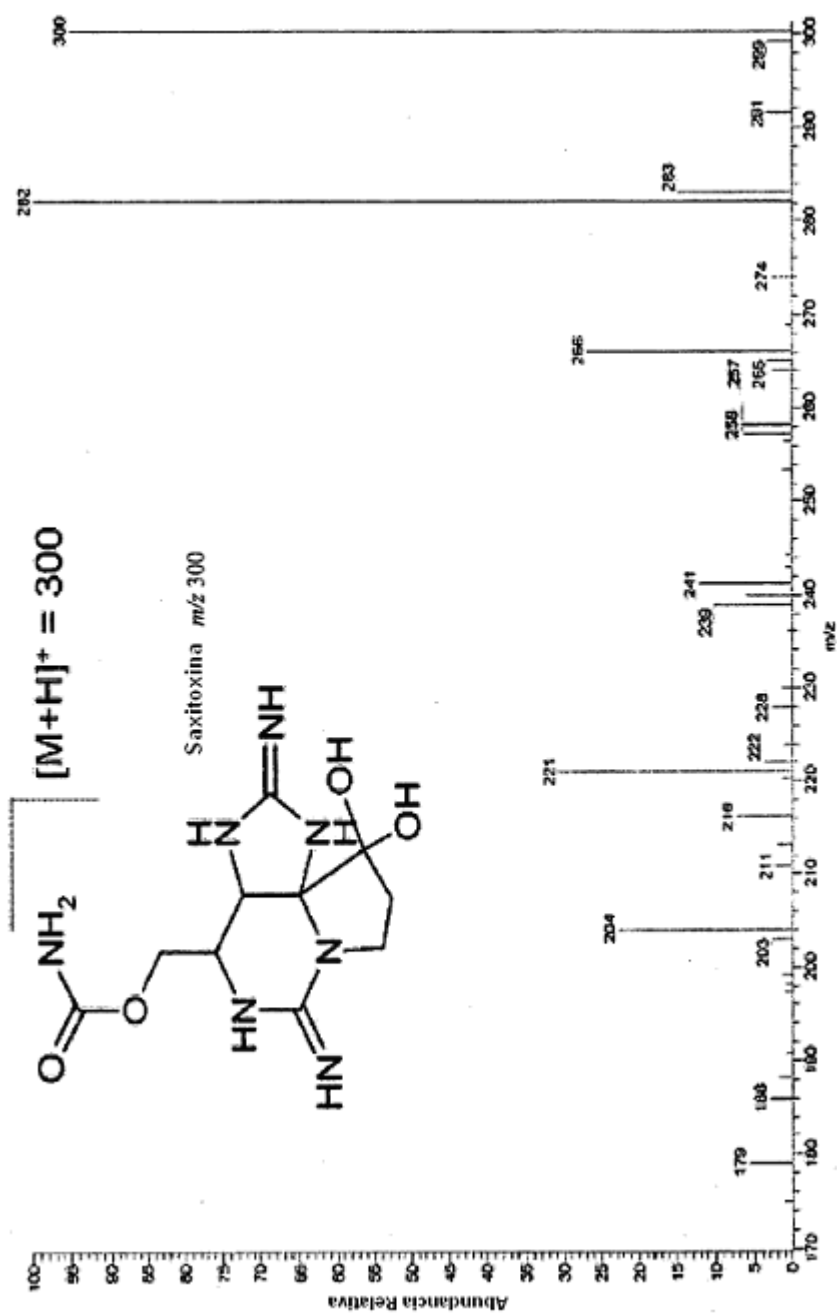


FIGURA 5B

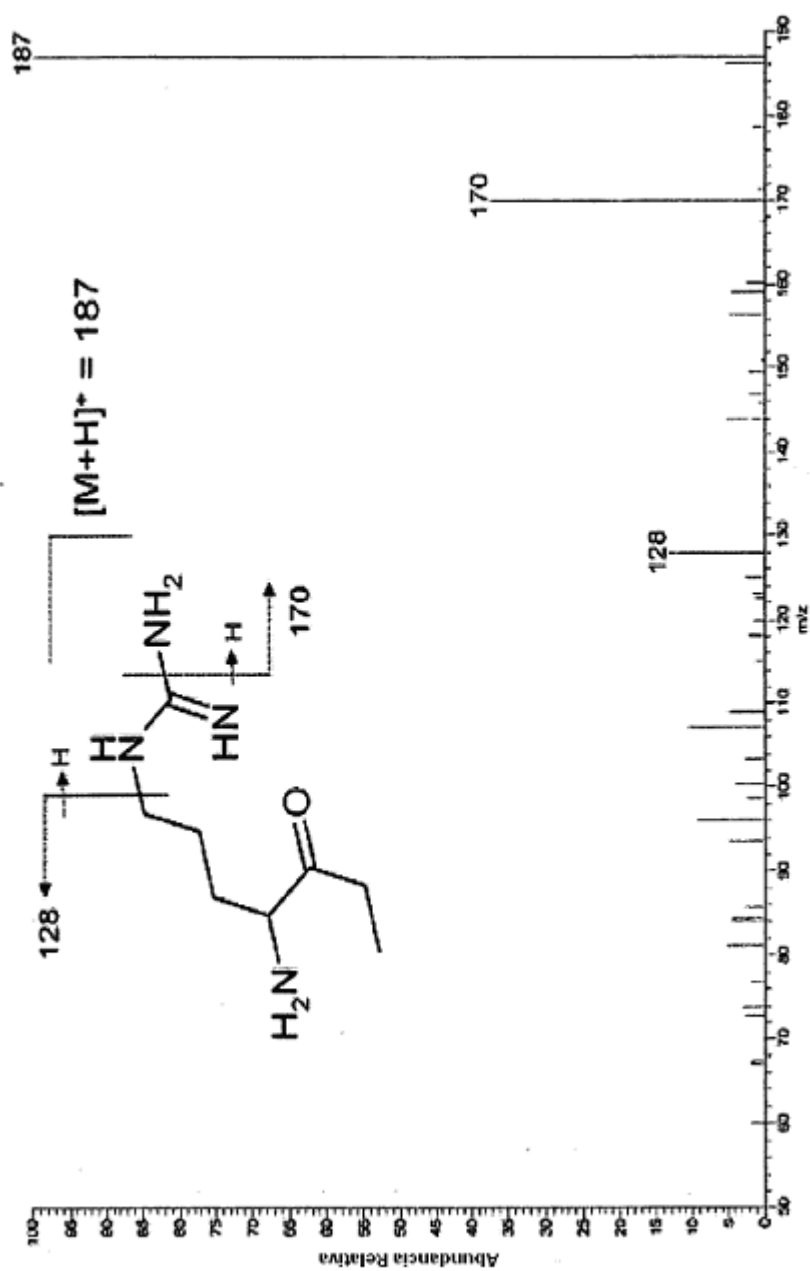


FIGURA 5C

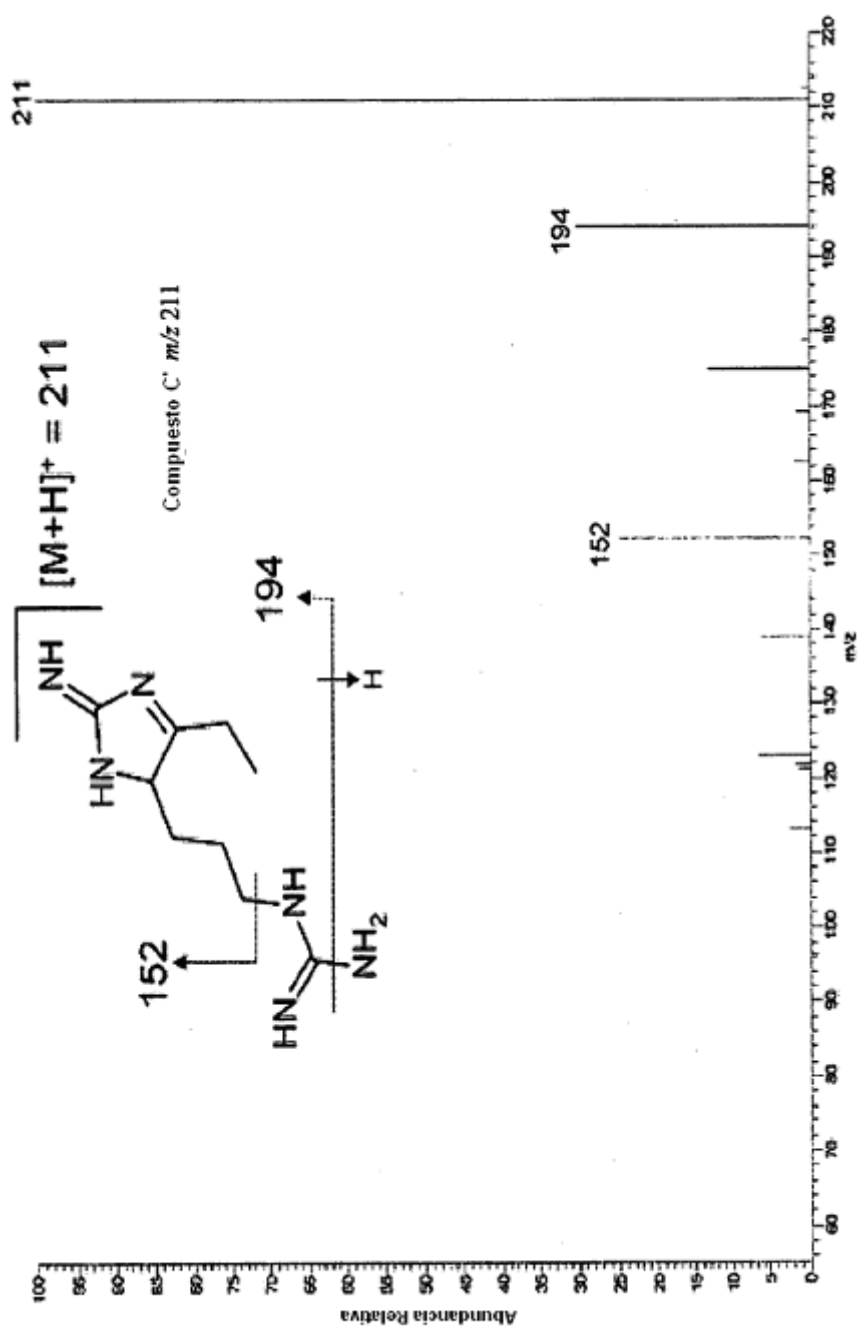


FIGURA 5D

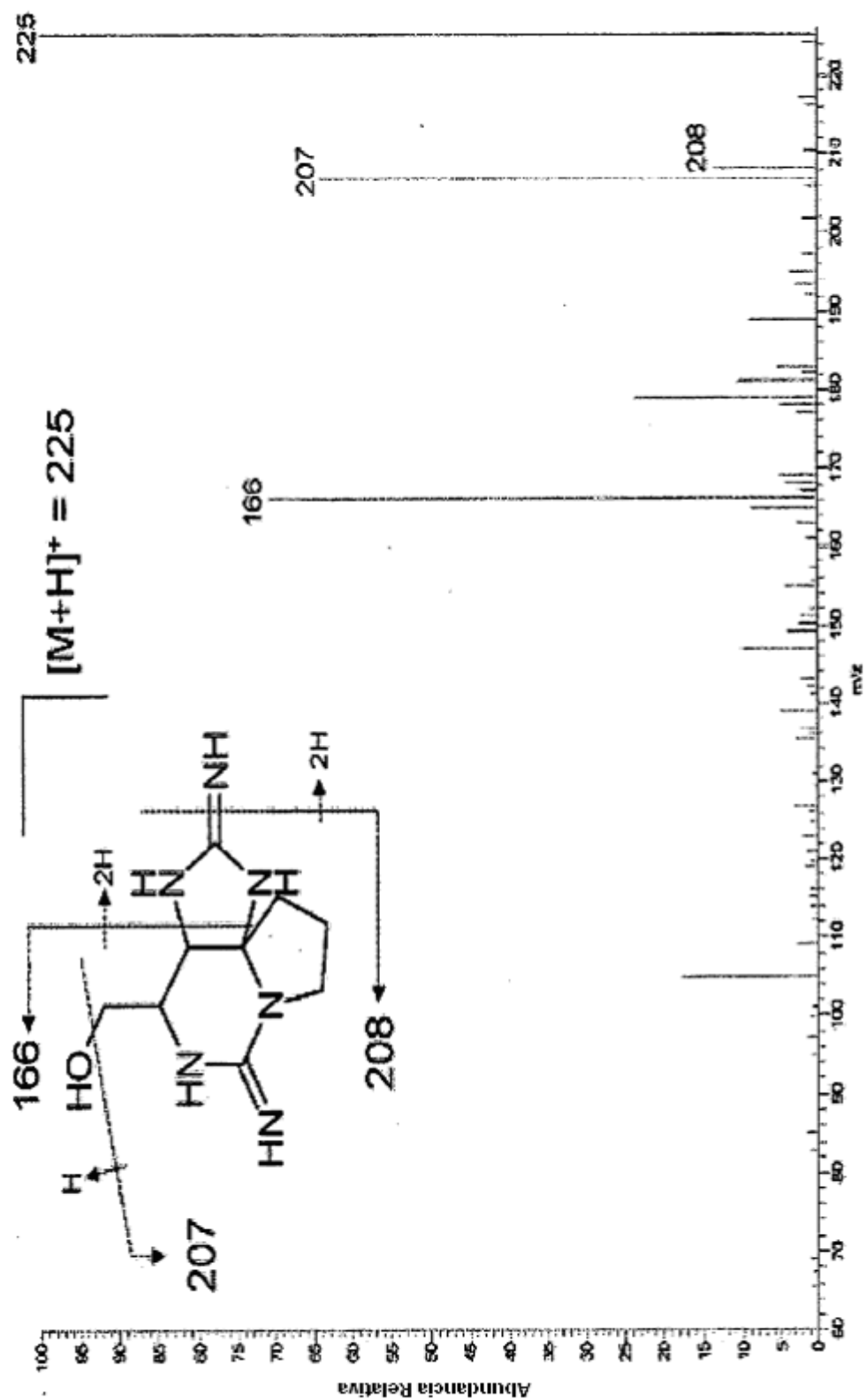


FIGURA 5E

Nombre	Familia de Enzimas	Tamaño (pb)	Concordancia de similitud Psi-Blast	% ID	Función Posible
<i>cyrD</i>	PKS CrpB	5631	ABM21570.1 <i>Nostoc sp.</i> ATCC 53789	58	PKS KS-AT-DH-KR-ACP
<i>cyrF</i>	PKS CrpB	4074	ABM21570.1 <i>Nostoc sp.</i> ATCC 53789	68	PKS KS-AT-ACP
<i>cyrG</i>	citosa desaminasa /Aminohidrolasa/ Dihidroorotasa	1437	BAF59909.1 <i>Pelotomaculum thermopropionicum</i> SI	50	Formación de anillo uracilo
<i>cyrI</i>	Proli 4-Hidroxilasa	831	ABB06365.1 <i>Burkholderia sp.</i> 383	43	Hidroxilación de C7
<i>cyrK</i>	Bomba de eflujo multifármaco MATE dirigida por Na ⁺	1398	EAW39051.1 <i>Lyngbya sp.</i> PCC 8106	65	Exportador
<i>cyrL</i>	Transposasa	750	ABG50981.1 <i>Trichodesmium erythraeum</i> IMS101	70	Transposasa
<i>cyrH</i>	citosa desaminasa /Aminohidrolasa/ Dihidroorotasa	1431	BAF59909.1 <i>Pelotomaculum thermopropionicum</i> SI	50	Formación de anillo uracilo
<i>cyrJ</i>	aminotransferasa de aminoácidos de cadena ramificada	780	<i>Trichodesmium erythraeum</i> IMS101	53	sulfotransferasa
<i>cyrA</i>	Amidinotransferasa AoaA	1176	AAX81898.1 <i>Cylindrospermopsis raciborskii</i>	100	Amidinotransferasa
<i>cyrB</i>	NRPS/PKS AoaB	8754	AAM33468.1 <i>Aphanizomenon ovalisporum</i>	97	NRPS/PKS A-dominio, pp, KS, AT, DH, Met, KR, ACP
<i>cyrE</i>	PKS	5667	ABA23591.1 <i>Anabaena variabilis</i> ATCC 29413	62	PKS KS-AT-DH-KR-ACP
<i>cyrC</i>	PKS AoaC	5005	AAM33470.1 <i>Aphanizomenon ovalisporum</i>	97	PKS KS-AT-KR-ACP
<i>cyrM</i>	Transposasa Parcial	318	ABG50981.1 <i>Trichodesmium erythraeum</i> IMS101	70	Transposasa
<i>cyrN</i>	Adenililsulfato quinasa (PAPS)	600	CAM76460.1 <i>Magnetospirillum gryphiswaldense</i> MSR-1	75	Adenililsulfato quinasa (PAPS)
<i>cyrO</i>	proteína hipotética	1548	EAW46978.1 <i>Nodularia spumigena</i> CCY9414	74	Regulador

FIGURA 6

Cepa de Cianobacteria	ARNr 16s	<i>cyrJ</i>	Toxicidad	Referencia
<i>Cylindrospermopsis raciborskii</i> T3	+	-	SXT	Lagos et al. (1999)
<i>Anabaena circinalis</i> 344B	+	-	N.D.	AWQC
<i>Cylindrospermopsis raciborskii</i> Germ1	+	-	N.D.	Neilan et al. (2003)
<i>Anabaena circinalis</i> 310F	+	-	N.D.	AWQC
<i>Cylindrospermopsis raciborskii</i> 44D	+	-	N.D.	NA
<i>Anabaena circinalis</i> 118C	+	-	SXT	Fergusson et al. (2000)
<i>Anabaena circinalis</i> 323A	+	-	N.D.	AWQC
<i>Anabaena circinalis</i> 323H	+	-	N.D.	AWQC
<i>Cylindrospermopsis raciborskii</i> VOLL2	+	-	N.D.	Neilan et al. (2003)
<i>Cylindrospermopsis raciborskii</i> VOLL1	+	-	N.D.	Neilan et al. (2003)
<i>Cylindrospermopsis raciborskii</i> HUNG1	+	-	N.D.	NA
<i>Cylindrospermopsis raciborskii</i> 023B	+	+	CYLN	Wilson et al. (2000)
<i>Cylindrospermopsis raciborskii</i> 05E	+	+	CYLN	Schembri et al. (2001)
<i>Cylindrospermopsis raciborskii</i> 4799	+	+	CYLN	Neilan et al. (2003)
<i>Cylindrospermopsis raciborskii</i> 24C	+	+	CYLN	Schembri et al. (2001)
<i>Cylindrospermopsis raciborskii</i> AWT 205	+	+	CYLN	Hawkins et al. (1997)
<i>Aphanizomenon ovalisporum</i> AO/QH	+	+	CYLN	NA

FIGURA 7

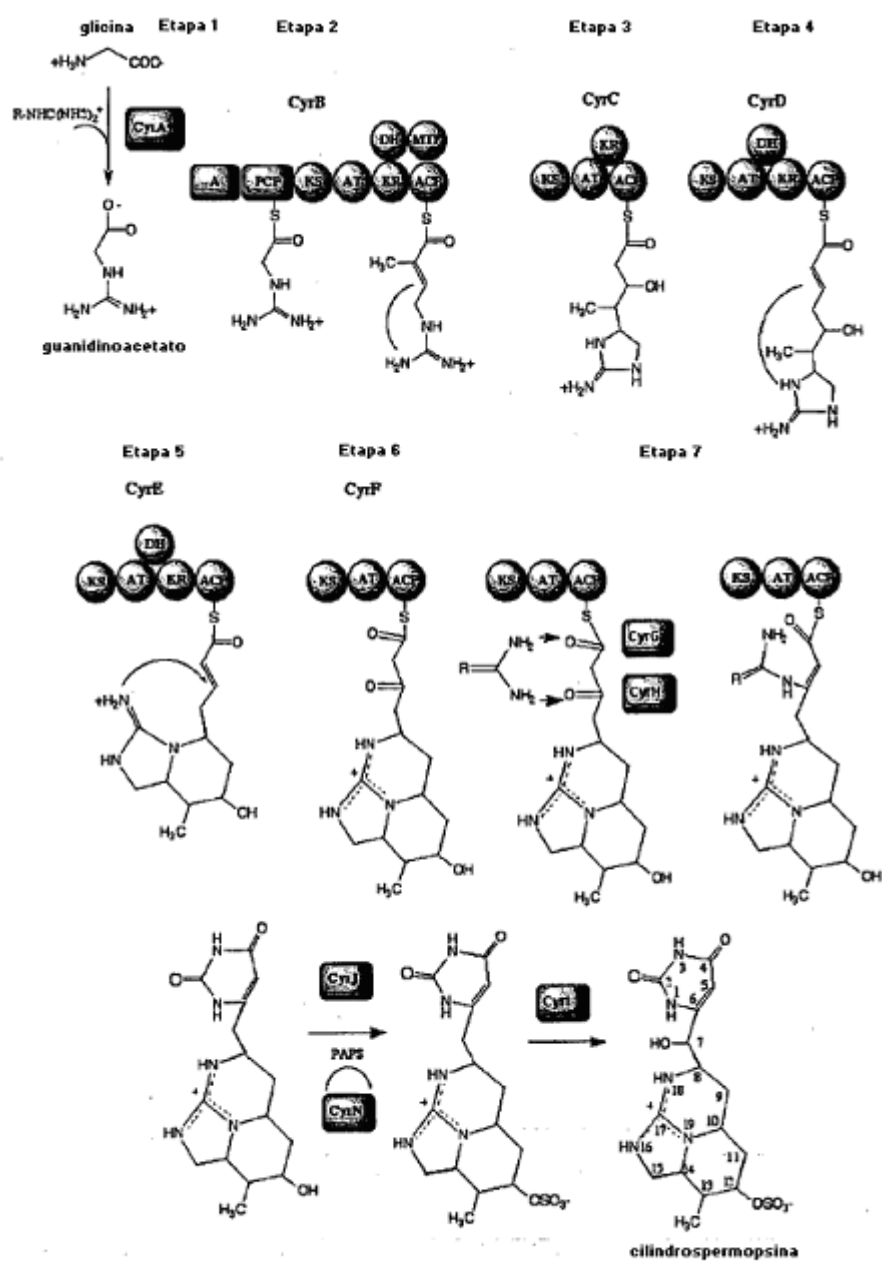


FIGURA 8

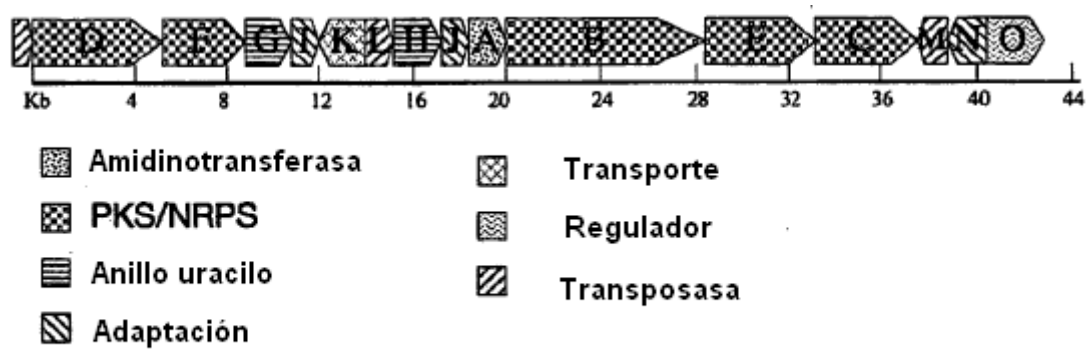


FIGURA 9