



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 298 276**

(51) Int. Cl.:

C12N 15/04 (2006.01)

C07K 14/40 (2006.01)

C12N 15/80 (2006.01)

C12N 1/15 (2006.01)

C12P 7/44 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **01984336 .6**

(86) Fecha de presentación : **25.07.2001**

(87) Número de publicación de la solicitud: **1303595**

(87) Fecha de publicación de la solicitud: **23.04.2003**

(54) Título: **Gen b5 de citicromo y proteína de *Candida tropicalis* y procedimientos correspondientes.**

(30) Prioridad: **26.07.2000 US 220958 P**
24.07.2001 US 912161

(45) Fecha de publicación de la mención BOPI:
16.05.2008

(45) Fecha de la publicación del folleto de la patente:
16.05.2008

(73) Titular/es: **Cognis IP Management GmbH**
Henkelstrasse 67
40589 Düsseldorf, DE

(72) Inventor/es: **Craft, David, L.;**
Madduri, Krishna, M. y
Loper, John, C.

(74) Agente: **Carvajal y Urquijo, Isabel**

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

Gen b5 de citocromo y proteína de *Candida tropicalis* y procedimientos correspondientes.

5 Antecedentes

Referencia cruzada a solicitudes relacionadas

Esta solicitud reivindica la prioridad para la solicitud provisional U.S. No. de serie 60/220, 958 presentada en Julio 26, 2000.

1. Campo de la invención

Esta invención se relaciona con procesos y composiciones involucrados en la producción de ácido dicarboxílico en levadura. Más particularmente, la invención se relaciona con un anfitrión todo transformado con un ácido nucleico que codifica una proteína b5 de citocromo.

2. Descripción de la técnica relacionada

Los ácidos alifáticos diácidos son intermedios químicos versátiles útiles como materia prima para la preparación de perfumes, polímeros, adhesivos y antibióticos macrolida. Mientras varias rutas químicas para la síntesis de ácidos α , ω -dicarboxílicos de cadena largas están disponibles, la síntesis no es fácil y la mayoría de métodos resulta en mezclas que contienen longitudes de cadenas más cortas. Como un resultado, son necesarias etapas de purificación extensas. Mientras se sabe que los ácidos diácidos de cadena larga también se pueden producir por transformación microbiana de alcanos, ácidos grasos o ésteres de estos, la síntesis química ha permanecido en la mayoría de la ruta disponible comercialmente, debido a limitaciones con los actuales enfoques biológicos.

Varias cepas de levadura se conocen por excretar ácidos α , ω -dicarboxílicos como un sub-producto cuando se cultiva sobre alcanos o ácidos grasos como la fuente de carbono. En particular, el almidón pertenece al género *Candida*, tal como *C. albicans*, *C. cloacae*, *C. guilliermondii*, *C. intermedia*, *C. lipolytica*, *C. maltosa*, *C. parapsilosis* y *C. zeylanoides* se conocen por producir tales ácidos dicarboxílicos (Agr. Biol. Chem. 35:2033-2042 (1971)). También, varias cepas de *C. tropicalis* se conocen por producir ácidos dicarboxílicos que varían en longitudes de cadena de C₁₁ a C₁₈ (Okino *et al.*, BM Lawrence, BD Mookherjee y BJ Willis (Eds), in flavors and fragrances: A World Perspective. Proceedings of the 10th International Conference of Essential oils, Flavors and Fragrances: Elsevier Science Publishers BV Amsterdam. (1988)), y son la base de varias patentes como se revisó por Bühler y Schindler, en Aliphatic Hydrocarbons in Biotechnology, H. J. Rehm and G. Reed (eds), vol. 169, Verlag Chemie, Weinheim (1984).

Estudios de los procesos bioquímicos mediante los cuales las levaduras metabolizan alcanos y ácidos grasos han revelado tres tipos de reacciones de oxidación: α -oxidación de alcanos a alcoholes, ω -oxidación de ácidos grasos a ácidos α , ω -dicarboxílicos y la β -oxidación degradante de ácidos grasos a CO₂ y agua. Los primeros dos tipos de oxidaciones se catalizan mediante enzimas microsómicas mientras que el último tipo tiene lugar en las peroxisomas. En *C. Tropicalis*, la primera etapa en la ruta de la ω -oxidación se cataliza mediante un complejo de enzima unido a membrana (complejo M-hidroxilasa) que incluye una monooxigenasa P450 citocromo y una reductasa citocromo NADPH. Este complejo hidroxilasa es responsable de la oxidación primaria del grupo metilo terminal en alcanos y ácidos grasos como se describe, por ejemplo, en Gilewicz *et al.*, Can. J. Microbiol 25:201 (1979). Los genes que codifican el citocromo P450 y los componentes reductasa NADPH del complejo se han identificado previamente como P450 ALK y P450 RED respectivamente, y también se han clonado y secuenciado como se describe, por ejemplo, en Sanglard *et al.*, Gene 76: 121-136 (1989). El P450ALK también ha sido designado P450ALK1. Más recientemente, los genes ALK se han designado por el símbolo CYP y los genes RED se han designado CPR. Ver, por ejemplo, Nelson, Pharmacogenetics 6(1). 1-42 (1996). Ver también Ohkuma *et al.*, DNA and Cell Biology 14:163-173 (1995). Seghezzi *et al.*, DNA and Cell Biology, 11:767-780 (1992) and Kargel *et al.*, Yeast 12:333-348 (1996). Por ejemplo, el P450ALK también se designa CYP52 de acuerdo con la nomenclatura de Nelson, *supra*. Los ácidos grasos se forman finalmente a partir de alcanos después de dos etapas adicionales de oxidación, se cataliza mediante oxidasa alcohol como se describe, por ejemplo, en Kemp *et al.*, Appl. Microbiol. And Biotechnol. 28:370-374 (1988), y deshidrogenasa aldehído. Los ácidos grasos se pueden oxidar adicionalmente a través de la misma ruta o ruta similar al ácido dicarboxílico correspondiente. La ω -oxidación de ácidos grasos procede por vía del ácido graso, ω -hidroxi y su derivado aldehído, con el ácido dicarboxílico correspondiente sin el requerimiento de la activación COA. Sin embargo, ambos ácidos grasos y ácidos dicarboxílicos se pueden degradar, después de activación con el correspondiente éster acil-COA a través de la ruta de β oxidación en las peroxisomas, que conduce a un acortamiento de cadena. En sistemas de mamíferos, los productos de ácido dicarboxílico y ácidos grasos de M-oxidación se activan con sus COA-ésteres en proporciones iguales y son sustratos para la β -oxidación peroxisómica y mitocondrial (J. Biochem 102: 225-234 (1987)). En levadura, la β -oxidación tiene lugar únicamente en las peroxisomas (Agr. Biol. Chem. 49:1821-1828 (1985)).

La producción de ácidos dicarboxílicos mediante fermentación de ácidos monocarboxílicos C14-C16 insaturados que utilizan una cepa de las especies *C. tropicalis* se describe en la patente U.S. 4,474,882. Los ácidos dicarboxílicos insaturados corresponden a los materiales de partida en el número y posición de los enlaces dobles. Procesos similares en los que se utilizan otros microorganismos especiales se describe en las patentes 3,975, 234 y 4,339,536, en la

descripción de patente Británica 1,405,026 y en las publicaciones de patente alemanas 21 64 626, 28 53 847, 29 37 292, 29 51 177 y 21 40 133.

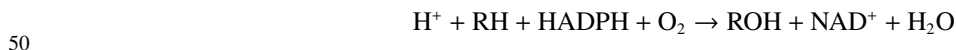
- La US 5,620,878 describe un proceso para producir ácidos α , ω -dicarboxílicos que comprenden cultivar una cepa
 5 *C. tropicalis* que tiene todos los cuatro genes POX alterados y que tienen un número de copia incrementado de por lo menos un gen P450ALK, por lo menos un gen P450RED, o una combinación de dichos genes en un medio de cultivo que contiene una fuente de nitrógeno, un sustrato orgánico y un co-sustrato.

- Las mono-oxigenasas P450 citocromo (P450S) son mono-oxidasas terminales de un sistema de enzima multi
 10 componente que incluye P450 y CPR. En algunos casos, un segundo portador de electrones, citocromo b5 (CYTb5) y su reductasa asociada están involucrados como se describe adelante en Morgan, *et al.*, Drug Metab. Dispensador. 12:358-364, 1984. El P450s comprende una super familia de proteínas que existe ampliamente en la naturaleza que han sido aisladas de una variedad de organismos como se describen en, Nelson, *Supra*. Estos organismos incluyen varios mamíferos, peces, invertebrados, plantas, moluscos, crustáceos, eucariotes inferiores y bacterias (Nelson, *Supra*). El
 15 primero se descubrió en microsomas de hígado de roedor como un pigmento de unión de monóxido de carbono como se describió, por ejemplo, en Garfinkel. Arch. Biochem. Biophys. 77:493-509 (1958), los P450s son denominados después con base en su absorción en 450 nm en un espectro de diferencia acoplado de CO reducido como se describe, por ejemplo, en Omura *et al.* Biol. Chem. 239:2370-2378 (1964).

- Los P450s catalizan el metabolismo de una variedad de compuestos endógenos y exógenos como se describió,
 20 por ejemplo, en Nelson, *supra*, y Nerbert *et al.*, *Supra*). Los compuestos endógenos incluyen esteroides, prostanoides, eicosanoides, vitaminas solubles en grasa, ácidos grasos, alcaloides de mamíferos, leucotrininas, aminos biogénicas y fitolexinas (Nelson, *supra* and Nebert *et al.*, *supra*). El metabolismo P450 involucra tales reacciones tales como epoxidación, hidroxilación, desalquilación, hidroxilación, sulfoxidación, desulfuración y deshalogenación reductora.
 25 Estas reacciones hacen generalmente que el compuesto sea más soluble, que sea conductor para excreción, y sea más electrófilo. Estos productos electrófilos pueden tener efectos dañinos si ellos reaccionan con el ADN u otros constituyentes celulares. Sin embargo, ellos pueden reaccionar a través de la conjugación de sustancias hidrófilas de bajo peso molecular que resultan en glucoronidación, sulfación, conjugación de aminoácidos o conjugación de glutatona que conduce típicamente a inactivación y eliminación como se describió, por ejemplo, en Klaassen *et al.*,
 30 toxicología, 3rd ed, Macmillan, New York, 14986.

- Los P450s son proteínas hemo tiolato que consiste de una porción hemo unida a una cadena de polipéptido sencilla de 45000 a 55000 da. El hierro del grupo hemo prostético se ubica en el centro de un anillo protoporfirino. Cuatro
 35 ligandos del hemo hierro se pueden atribuir al anillo porfirino. El quinto ligando es un anión de tiolato de un residuo cisteinilo del polipéptido. El sexto ligando es probablemente un grupo hidroxilo de un residuo de aminoácido, o una porción con una resistencia de campo similar tal como una molécula de agua como se describe, por ejemplo, en Goeptar *et al.*, Critical Reviews in toxicology 25(1):25-65 (1995).

- Las reacciones de mono-oxigenación catalizadas por citocromos P450 en un sistema de membrana unida eucariótico requiere la transferencia de electrones de NADPH a P450 por vía de reductasa P450 NADPH-citocromo (CPR) como se describe, por ejemplo, en Taniguchi *et al.*, Arch. Biochem. Biophys. 232:585 (1984). Los genes CPR se refieren ahora como genes NSP. Ver, por ejemplo, Debacker *et al.*, Antimicrobial Agents and Chemotherapy, 45:1660 (2001). El CPR es una flavoproteína de aproximadamente 78000 Da que contiene 1 mol de dinucleótido adenina flavina (FAD) y 1 mol de mono-nucleótido de flavina (FMN) por mol de enzima como se describe, por ejemplo, en Potter
 45 *et al.*, J. Biol. Chem, 258:6906 (1983). La porción FAD de CPR es el sitio de entrada de electrones en la enzima, mientras el FMN es el sitio donante de electrones para el P450 como se describe, por ejemplo, en Vernillon *et al.*, J. Biol. Chem. 253:8812 (1978). La reacción general es como sigue:



- Posterior a la transferencia del primer electrón, el O_2 se une al complejo del sustrato Fe_2^+ -P450 para formar el complejo del sustrato Fe_3^+ -P450. Este complejo se reduce luego mediante un segundo electrón del CPR, o, en algunos casos, NADH por vía de un segundo portador de electrones, b5 citocromo (CYTb5) y su reductasa b5 citocromo
 55 NADH asociada como se describe, por ejemplo, en Guengerich *et al.*, Arch. Biochem. Biophys. 205:365 (1980), and Morgan, *supra*. La mayoría de estudios mencionados anteriormente implican al CYTb2 por estar involucrado en la ruta solo para la transferencia del segundo electrón. Un átomo de este oxígeno reactivo se introduce en el sustrato, mientras que el otro se reduce a agua. El sustrato oxigenado se disocia luego, regenerando la forma oxidada del P450 citocromo como se describe, por ejemplo, en Klassen., Amdur and Doull, Casarett and Doull's Toxicology, Macmillan,
 60 New York (1986).

- Con respecto al CYTb5, se han propuesto otros varios modelos del papel de esta proteína en la expresión P450 a pesar de su papel como un portador de electrones. Otro modelo del papel CYTb5 en la expresión P450 se basa en los efectos de las interacciones de proteína como se describe, por ejemplo, en Tamburini *et al.*, Proc. Natl. Acad. Sci.
 65 8A:11-15 (1986). Mediante este modelo, la unión del CYTb5 con P450 resulta en un complejo con alto contenido incrementado de espín. Este complejo luego tiene una mayor afinidad para el CPR. A través de esta interacción, el CYTb5 no proporciona actualmente el segundo electrón, pero reduce el tiempo entre la primera y segunda transferencia de electrones. Estas conclusiones se hacen en base en el trabajo con conejos CYP2b4. Adicionalmente, el CYTb5 no

evita la interacción del conejo CYP2b4 con CPR. Por lo tanto, se propone que los dominios de unión de P450 para el CYTb5 y el CPR sena diferentes. Esto puede explicar posiblemente porque solo ciertos OP450 interactúan con CYTb5. Se sabe que el CYTb5 se puede reducir mediante reductasa b5 citocromo o CPR como se describe, por ejemplo, en Enoch *et al.*, J. Biol. Chem, 254:8976-8981, (1979). Sin embargo, en por lo menos un caso interesante, la adición de la reductasa b5 citocromo a un sistema reconstituido que contiene CYTb5 no resulta en una oxidación de sustrato incrementada como se describe, por ejemplo, en Sugiyama *et al.*, J. Biochem. 87:1457-1467 (1979).

Varios otros estudios también han indicado que el CYTb5 puede regular la fosforilación del P450 citocromo mediante varias proteínas quinasa, por ejemplo, la quinasa proteína dependiente de cAMP y la quinasa proteína C como se describe, por ejemplo, en Jansson *et al.*, Arch. Biochem. Biophys. 259:441-448 (1987), Epstein *et al.*, Arch. Biochem. Biophys. 271(2):424-432(1989) and Lobanov *et al.* Biokhimiia 58(10):1529-1537 (1993).

Independiente el mecanismo, numerosos estudios indican que el involucramiento del CYTb5 es ventajoso en algunas reacciones mediadas por P450. Un papel obligatorio del CYTb5 se ha mostrado en varios sistemas P450 reconstituidos como se describe, por ejemplo, en Sugiyama *et al.*, *supra*, Sugiyama *et al.* Biochem. and Biophys. Res. Comm., 90:715-720 (1979), Canova-Davis *et al.*, J. Biol. Chem. 259:2541-254b,1983, Kuwahara *et al.*, Biochem. and Biophys. Res. Comm., 96:1562-1568 (1980), and Sasame *et al.*, Life Sciences 14:35-46 (1974). En todos estos casos, la actividad es abolida luego de la omisión del CYTb5 agregado exógenamente o su inactivación mediante la adición del anticuerpo CYTb5. Como se anotó anteriormente, el CYTb5 también está involucrado en la ruta del segundo electrón en ciertas reacciones mediadas por P450 como se describe, por ejemplo, en Hrycay *et al.*, Archv. Bioch. 165:331-339 (1974), Imai *et al.*, Biochem. and Biophys. Res. Comm. 75:420-426 (1977), and Imai, J. Biochem. 89:351-362 (1980), así como también la reducción del oxi-citocromo P450 para el complejo de oxígeno como se describe, por ejemplo, en Noshiro *et al.*, J. Biochem. 116:521-526 (1981).

En otros estudios, que utilizan sistemas reconstituidos, el CYTb5 agregado exógenamente incrementa la actividad del CYP2B4 de conejo como se describe, por ejemplo, en Sugiyama *et al.*, J. Biochem., 92:1793-1803 (1982), and Chiang, Archv. Bioch. Biophy., 211:662-673 (1981), como se describe, por ejemplo, en Vatsis *et al.*, J. Biol. Chem. 257:11221-11229 (1982), perros CYP2B11 y ratas CYP2B1 como se describe, por ejemplo, en Duignan *et al.*, Arch. Bioch. Biophy. 267:294-304 (1988), y una reacción de hidroxilación de clorobenceno de rata P450 como se describe, por ejemplo, en Lu *et al.*, Biochem. and Biophys. Res. Comm. 61:1348:1355 (1974).

El CYP51, desmetilasa 14 α lanosterol, desarrollan una reacción esencial en la biosíntesis del ergosterol, la mayor membrana esteroil de *Saccharomyces cerevisiae* (Sc) como se describe, por ejemplo, en Parks, CRC Crit. Rev. Microbiol., 6:301-341 (1978). Las cepas Sc que sintetizan el ergosterol no utilizan ergosterol, sino que aquellas se bloquean en las etapas pre-esterol o en la hemo biosíntesis que puede tomar y utilizar esteroil exógeno como se describe, por ejemplo, en Parks, *supra*. El crecimiento anaeróbico del Sc no puede sintetizar esteroides pero utilizan ergosterol exógeno. La interrupción del CYP51 produce cepas Sc que continúan produciendo esteroides 14 α -metilo, no producen ergosterol detectable y son incapaces del crecimiento aeróbico como se describe, por ejemplo, en Kalb *et al.*, DNA, 6:529-537 (1987). A diferencia ellos son anaerobios estrictos, una condición en donde ellos acumulan ergosterol exógeno.

El CPR funciona como el donante de electrones para su reacción desmetilasa como se describe, por ejemplo, en Aoyama *et al.*, J. Biol. Chem. 259:1661-1666 (1984). Si el CPR es el único donante de electrones de CYP51 en Sc, mutantes completos CPR (cpr1) deben semejar fenotípicamente las cepas CYP51 incompletas (cyp51). Sin embargo, un mutante completo cpr1 Sc no es estrictamente anaeróbico, y produce ergosterol como se describe, por ejemplo, en Sutter and Loper, Biochem. and Biophys. Res. Comm., 160:1257-1266 (1989). La producción de ergosterol muestra que algunos CYP51 son todavía funcionales. Los genes responsables para esta recuperación de un mutante completo cpr1 se muestra por ser aquel que codifica el CYTb5. lo que muestra que en este sistema, el CYTb5 es capaz de imitar el CPR funcionalmente cuando está presente en un alto número de copia como se describe, por ejemplo, en Truan *et al.*, Gene, 142(1):123-7 (1994).

La expresión del mamífero P450s en Sc se ha logrado con resultados variantes como se describe, por ejemplo, en Renaud *et al.*, J. Biochem. 194:889-896 (1990), Urban *et al.*, Biochimie 72:463-472 (1990), and Pompon *et al.*, Molecular Endocrinology 3:1477-1487 (1989). Estudios han mostrado que la actividad por P450s de varios mamíferos expresado en Sc han sido mejorados luego de la adición de CYTb5 para aislar microsomas de estas cepas. Específicamente, la actividad Cyp1a-1 de ratón y CYP3A4 humana microsómica Sc se incrementan luego de la adición de CYTb5 de conejo como se describe, por ejemplo, en Renaud *et al.*, *supra*, and Urban *et al.*, *supra*.

Adicionalmente a sus efectos positivos, un estudio que involucra el Sc sugiere que el CYTb5 puede actuar negativamente como se describe, por ejemplo, en Pompon *et al.*, *supra*. La adición de CYTb5 de conejo para aislar microsomas de un CYP19 humano que expresa Sc resulta en una reducción en la actividad aromatasa.

Los ácidos dicarboxílicos alifáticos ($\leq$ C12) de cadena corta (diácidos) son intermedios industriales importantes en la fabricación de diésteres y polímeros, y encuentran aplicación como termoplásticos, agentes plastificantes, lubricantes, fluidos hidráulicos, químicos agrícolas, farmacéuticos, tintes, tensoactivos, y adhesivos. El alto precio y disponibilidad limitada de los diácidos de cadena corta se debe a las restricciones impuestas por la síntesis química existente.

Los diácidos de cadena larga (ácidos α , ω -dicarboxílicos alifáticos con números de carbono de 12 o más, también referidos de aquí en adelante como diácidos) ($\text{HOOC}-(\text{CH}_2)_n-\text{COOH}$) son una familia versátil de químicos con utilidad potencial y demostrada en una variedad de productos químicos que incluyen plásticos, adhesivos, y fragancias. Desafortunadamente, el mercado potencial completo de diácidos no se ha realizado debido a que los procesos químicos producen solo un rango limitado de estos materiales en un precio relativamente alto. Adicionalmente, los procesos químicos para la producción de diácidos tienen un número de limitaciones y desventajas. Todos los procesos químicos se restringen a la producción de diácidos de longitud de cadena de cadenas de carbono específicas. Por ejemplo, el proceso del ácido dodecanodióico inicia con butadieno. Los diácidos de producto resultante se limitan a múltiples longitudes de 4 carbonos y, en la práctica, solo se hace el ácido dodecanóico. Los procesos dodecanóicos se basan en materia prima petroquímica no renovable. Los procesos de conversión multi reacción producen sub-productos indeseados, que pueden resultar en pérdidas de producto, polución NO_x y desprecios de metales pesados.

Los diácidos de cadena larga ofrecen desventajas potenciales sobre los diácidos de cadenas mas cortos, pero su alto precio de venta y disponibilidad comercial limitada evitan el creciente desperdicio en muchas de estas aplicaciones. Los biocatalizadores ofrecen una forma innovadora de superar estas limitaciones con un proceso que produce un amplio rango de productos diácidos de materias primas renovables. Sin embargo, no existen bioprocesos viables comercialmente para producir diácidos de cadena larga a partir de recursos renovables.

Resumen de la invención

De acuerdo con la presente invención, se proporciona una célula anfitriona transformada con un ácido nucleico que codifica la proteína CYTb5 (SEQ. ID. NO. 2), en donde todos los anfitriones contienen una interrupción de la ruta de β -oxidación.

Un método para incrementar la producción de ácido dicarboxílico se proporciona el cual incluye suministrar una célula anfitriona que tiene ninguno, uno o más genes CYTb5; que incrementa, en la célula anfitriona, el número de genes CYTb5 que codifica una proteína CYTb5 que tiene la secuencia de aminoácido establecida en la SEQ. ID NO. 2 en donde todos los anfitriones contienen una interrupción de la ruta de β -oxidación; y cultivar la célula anfitriona en medio que contiene un sustrato orgánico que induce la expresión del gen CYTb5, para efectuar la producción incrementada de ácido dicarboxílico.

Breve descripción de los dibujos

Las figuras 1A-1C describe la secuencia de nucleótido del gen CYTb5 (SEQ. ID. NO. 1) y la secuencia de aminoácido de la proteína CYTb5 (SEQ. ID. NO. 2).

Las figuras 2A-2B describen la secuencia de nucleótido del gen CPRA (SEQ. ID. NO. 3).

Las figuras 3A-3C describen la secuencia de aminoácido de la proteína CPRA (SEQ. ID. NO. 4).

Las figuras 4A-4B describen la secuencia de nucleótido del gen CPRB (SEQ. ID. NO. 5).

Las figuras 5A-5C describen la secuencia de aminoácido de la proteína CPRB (SEQ. ID. NO. 6).

Las figuras 6A-6C describen la secuencia de nucleótido del gen CYP52A1A (SEQ. ID. NO. 7).

Las figuras 7A-7B describen la secuencia de nucleótido del gen CYP52A2A (SEQ. ID. NO. 8).

Las figuras 8A-8C describen la secuencia de nucleótido del gen CYP52A2B (SEQ. ID. NO. 9).

Las figuras 9A-9C describen la secuencia de nucleótido del gen CYP52A3A (SEQ. ID. NO. 10).

Las figuras 10A-10C describen la secuencia de nucleótido del gen CYP52A3B (SEQ. ID. NO. 11).

Las figuras 11A-11C describen la secuencia de nucleótido del gen CYP52D4A (SEQ. ID. NO. 12).

Las figuras 12A-12C describen la secuencia de nucleótido del gen CYP52A5A (SEQ. ID. NO. 13).

Las figuras 13A-13C describen la secuencia de nucleótido del gen CYP52A5B (SEQ. ID. NO. 14).

Las figuras 14A-14B describen la secuencia de nucleótido del gen CYP52A8A (SEQ. ID. NO. 15).

Las figuras 15A-15C describen la secuencia de nucleótido del gen CYP52A8B (SEQ. ID. NO. 16).

La figura 16 es una representación esquemática del vector de clonación pTrip1Ex disponible de Clontech™ Laboratories, Inc. Se muestran Los sitios de restricción seleccionados dentro del sitio de clonación múltiple.

ES 2 298 276 T3

La figura 17A es un mapa del vector ZAP Express™.

La figura 17B es una representación esquemática de un fagémido de clonación, vector pBK-CMV.

5 La figura 18 es una descripción esquemática de un procedimiento para extraer y analizar diácidos y monoácidos a partir de cultivos de fermentación.

10 La figura 19 es una representación esquemática del plásmido pCR2.1™ disponible de Invitrogen. Las secuencias de ácido nucleico para los sitios de restricción seleccionados y otras características se describen (SEQ. ID. NO. 43 y hebra complementaria SEQ. ID. NO. 44).

La figura 20 es una representación esquemática del vector de expresión de *S. cerevisiae* pYES 2.0 disponible de Invitrogen. Se indican los sitios de restricción seleccionados.

15 La figura 21 es una representación esquemática de ciertas porciones conservadas de ciertos genes CPR.

La figura 22 es una representación esquemática del plásmido pHKM1. Los sitios de restricción seleccionados se indican así como también la posición del gen CPRA.

20 La figura 23 es una representación esquemática del plásmido pHKM4. Los sitios de restricción seleccionados se indican así como también la posición del gen CPRA.

La figura 24 es una representación esquemática del plásmido pHKM9. Los sitios de restricción seleccionados se indican así como también la posición del gen CPRB.

25 La figura 25 es una descripción gráfica de la proporción Log del producto desconocido (u) al competidor (c) versus la concentración de competidor de ARN en las reacciones QC-RT-PCR que se relacionan con CYTb5.

30 La figura 26 es una descripción esquemática de un procedimiento para regenerar un genotipo URA3A por vía de la recombinación homóloga que involucra los genes CYTb5 y/o CPR.

La figura 27 es una descripción esquemática de la recombinación homóloga que involucra una división del gen URA3A y la incorporación de los genes CPR y/o CYTb5.

35 La figura 28 describe la secuencia de nucleótidos del gen URA3A. (SEQ. ID. NO. 45).

La figura 29 es una descripción esquemática del plásmido pNEB193 disponible de New England Biolabs. Se indican los sitios de restricción seleccionados.

40 La figura 30 es una descripción esquemática del vector de integración pURAIin que contiene URA3A modificado en el plásmido pNEB193.

Descripción de las modalidades preferidas

45 La productividad diácido se mejora de acuerdo con la presente invención al incrementar selectivamente una proteína que es importante para la oxidación de sustratos orgánicos tal como ácidos grasos que componen el alimento deseado. Un gen CYTb5 único de *C. tropicalis* 20336 se ha identificado y caracterizado que está involucrado en las reacciones P450.

50 La secuencia de nucleótido (2710bp) del gen CYTb5 de longitud completa incluye una región reguladora y una región de codificación de proteína definida por nucleótidos 1109-1495 como se muestra en LAS Figuras 1A-1C. La secuencia de aminoácido (129 aminoácidos) de la proteína CYTb5 también se muestra en LAS Figuras 1A-1C.

55 La amplificación del gen CYTb5 en células anfitrionas, por ejemplo, levadura, que excreta ácidos dicarboxílicos, particularmente ácidos α , ω -dicarboxílicos. La expresión del gen CYTb5 de *C. tropicalis* y su inducción por alcanos o ácidos grasos se puede medir en células aisladas durante el curso de bioconversiones de fermentación utilizándole ensayo de reacción de cadena de polimerasa de transcripción inversa competitivo cuantitativo (QC RT-PCR) como se describe adelante (ejemplos 11 y 14).

60 Adicionalmente, la modificación de los promotores existentes y/o el aislamiento de promotores alternos proporciona expresión incrementada del gen CYTb5. Los promotores fuertes se obtienen de por lo menos cuatro fuentes: modificaciones específicas o aleatorias del promotor CYTb5, la selección de un promotor fuerte de los genes de β -oxidación de *Cándida* tal como POX4 y POX5, o la selección de otro promotor *Cándida* adecuado.

65 La fuerza del promotor se mide utilizando QT-RT-PCR para medir la expresión del gen CYTb5 en levadura, por ejemplo, células *Cándida*, aisladas de fermentadores. Los ensayos enzimáticos y los anticuerpos específicos para la proteína CYTb5 se utilizan para verificar que la fuerza del promotor incrementada se refleja mediante la síntesis incrementada de la proteína correspondiente. Una vez se ha identificado un promotor adecuado, este se fusiona para

seleccionar un gen CYTb5 e introducirlo en *Cándida* para la construcción de una nueva cepa producida mejorada. Se contempla que la región de codificación del gen CYTb5 se puede fusionar a promotores adecuados u otras secuencias reguladoras que se conocen por aquellos expertos en la técnica.

- 5 La productividad diácida se mejora así mediante la integración selectiva, amplificación, y sobre expresión del gen CYTb5 en el anfitrión de producción de *C. tropicalis*.

Se debe entender que las células anfitrionas dentro de las cuales una o más copias del gen CYTb5 deseado se han introducido se pueden hacer para incluir tal un gen median te una técnica conocida por aquellos expertos en el arte. Por ejemplo, las células anfitrionas adecuadas incluyen procariotes tal como *Bacillus sp.*, *Pseudomous sp.*, *Actinomy-*
 10 *cetes sp.*, *Eschericia sp.*, *Mycobacterium sp.*, y eucariotes tales como levadura, algas, células de insectos, células de plantas y hongos filamentosos. Las células anfitrionas adecuadas son células de levadura preferiblemente tal como *Yarrowia*, *Bebaromyces*, *Saccharomyces*, *Schizosaccharomyces*, *Pichia*, *Lipomyces*, *Rhodospiridium*, *Rhodotorula*, *Trichosporan* *Cryptococcus*, *Endomyces*, *Galactomyces*, *Williopsis*, *Waltomyces*, y más preferiblemente aquellas del
 15 género *Cándida*. Especies preferidas de *Cándida* son *tropicalis*, *maltosa*, *apicola*, *paratropicalis*, *albicans*, *cloacae*, *guillermundii*, *intermedia*, *lipolitica*, *parapsilosis* y *zeilenoides*. Particularmente los anfitriones preferidos incluyen cepas *C. tropicalis* que se han modificado genéticamente de tal manera que uno o más del POX4A cromosómico, POX4B y ambos genes POX5 se han interrumpido como se describe por ejemplo en la patente U.S. Nos. 5,254,466 y 5,620,878. Las interrupciones del gen POX4 y POX5 bloquean efectivamente la ruta β -oxidación en su primera reacción (que se cataliza mediante oxidasa acil-CoA) en una cepa de anfitrión *C. tropicalis*. Los genes POX4A y POX5 codifican
 20 distintas subunidades de oxidasa acil-CoA de cadena larga que son los polipéptidos peroxisomales (PXPs) designados PXP-4 y PXP-5, respectivamente. La interrupción de uno o más de estos genes resulta en una inactivación parcial o completa de la ruta de β -oxidación permitiendo así rendimientos mejorados del ácido dicarboxílico al redirigir el sustrato hacia la ruta ω -oxidación y también evita la reutilización de productos de ácido dicarboxílico a través de la
 25 ruta de β -oxidación.

Ejemplos de cepas de *C. tropicalis* que son beta-oxidación parcialmente bloqueados incluyen, H41, H41B, H51, H45, H43, H53, H534, H534B y H435 como se describe en la Patente U.S. mencionada anteriormente No. 5,254,466. Un Ejemplo de una cepa de beta-oxidación completamente bloqueada de *C. tropicalis* en donde todos los cuatro genes
 30 POX4 y POX5 se interrumpen por el marcador seleccionable URA3, es H5343 (ATCC 20962) como se describe en la patente U.S. No. 5,254,466.

Vectores tales como plásmidos, fagémidos, fagos, cósmicos, cromosomas de levadura artificial, plásmidos de levadura episódicos, plásmidos de levadura replicados, y similares se pueden utilizar para transformar o transfectar
 35 adecuadamente células anfitrionas. Las células anfitrionas también se pueden transformar al introducirlas en una célula de vector de ADN lineal que contienen la secuencia de gen deseado, es decir CYTb5. Tal ADN lineal puede ser ventajoso cuando es deseable evitar la introducción de ADN no nativo (externo) dentro de la célula. Por ejemplo, el ADN que consiste de CYTb5 flanqueado por secuencias de ADN que son nativas a la célula originaria se pueden introducir en la célula anfitriona por electroporación, transformación de acetato de litio, esferoplastia y similares.
 40 Las secuencias de ADN pueden incluir marcadores seleccionables y/o otras herramientas para ingeniería genética. Las células se levadura se pueden transformar con cualquiera de los vectores de expresión descritos aquí. El término “vector de expresión” se utiliza ampliamente aquí y está destinado a abarcar cualquier medio que puede ser utilizado para transformar una célula objetivo. El vector de expresión abarca todos los ejemplos de vectores listados aquí, que incluyen, por ejemplo, vectores de integración.

En una modalidad particularmente útil un vector que incluye el gen CYTb5 también puede incluir una o más copias de un gen CPR (reductasa P450 NADPH-citocromo). El gen CPR de la cepa de levadura, por ejemplo, *C. tropicalis*, que codifica el componente reductasa P450 NADPH-citocromo del complejo ω -hidroxilasa se ha clonado y secuenciado como se describe, por ejemplo, en Sanglard *et al.*, *supra*. Las cepas *C. tropicalis* que tienen un mayor
 50 número de genes CPR que la cepa tipo silvestre han mostrado productividad incrementada de ácidos dicarboxílicos como se describe, por ejemplo, en la patente U.S. mencionada anteriormente No. 5,620,878.

Ejemplos específicos de genes CPR incluyen los genes CPRA y CPRB de *C. tropicalis* 20336 como se describe, por ejemplo, en la patente U.S. No. 6,331,420 y WO 00/20566 A2. Las regiones de codificación y reguladoras completas
 55 para el gen CPRA de *C. tropicalis* 20336 y su secuencia de aminoácidos se muestran en LAS Figuras 2A-2B y 3A-3C, respectivamente. Las regiones de codificación y regulación completas para el gen CPRB de *C. tropicalis* 20336 y su secuencia de aminoácidos se muestra en LAS Figuras 4A-4B y 5A-5C, respectivamente. Se debe notar que hay alguna evidencia que, en *C. tropicalis*, el codón CTG no se traduzca como leucina de acuerdo con el “código genético universal”, sino como serina. Ver, por ejemplo, Ueda *et al.*, *Biochemie* (1994) 76, 1217-1222. Sin embargo, esta proposición no se ha probado de forma concluyente. De acuerdo con esto, en razón a que el codón CTG en la posición 1180-1182 de La figura 4A se puede traducir como una leucina o una serina, el aminoácido 50 mostrado en La figura 5A se designa “X” en donde “X” puede ser leucina o serina, de la misma manera, en razón a que el codón CTG en la posición 1153-1156 de La Figura, 2A se puede traducir como leucina o serina, el aminoácido 50 mostrado en La figura 3A se designa “X”, en donde “X” puede ser leucina o serina. Se debe mantener en mente que otras secuencias
 65 de ADN que incorporan el codón CTG que codifica una proteína como se describe aquí, puede tener similitud si aún no se designa con una “X”.

Un ejemplo de una cepa *C. tropicalis* que contiene copias adicionales de los genes CYTb5 y CPR de *C. tropicalis* 20336 es la cepa HDC11 (ver Tabla 2, Ejemplos 2 y 15).

El vector mencionado anteriormente incluye el gen CYTb5 que también puede incluir una o más copias de un gen CYP (monooxigenasa P450 citocromo). Genes CYP adecuados de levadura, por ejemplo, la cepa *C. tropicalis* 20336, incluye, pero no se limita a, CYP52A1A (LAS Figuras 6A-6C), CYP52A2A (LAS Figuras 7A-7B), CYP52A2B (LAS Figuras 8A-8C), CYP52A3A (LAS Figuras 9A-9C), CYP52A3B (LAS Figuras 10A-10C), CYP52D4A (LAS Figuras 11A-11C), CYP52A5A (LAS Figuras 12A-12C), CYP52A5B (LAS Figuras 13A-13C), CYP52A8A (LAS Figuras 14A-14B) y CYP52A8B (LAS Figuras 15A-15C), como se describe en la solicitud U.S. No. de serie No. 09/302,620 y la solicitud internacional No. PCT/US9920797.

Se proporciona un método para incrementar la producción de una proteína CYTb5 en una célula anfitriona. El método incluye (a) transformar la célula anfitriona que tiene un número de ocurrencia natural de los genes CYTb5 con por lo menos un gen CYTb5 que codifica una proteína CYTb5 que tiene la secuencia de aminoácido como se establece en LAS Figuras 1A-1C, y (b) cultivar la célula anfitriona transformada y por lo tanto incrementar la expresión de la proteína CYTb5 comparada con aquella de la célula anfitriona que contiene el número de la copia de ocurrencia natural de los genes CYTb5. Las células anfitrionas adecuadas incluyen aquellas descritas anteriormente. Preferiblemente, la célula anfitriona es una célula cándida, y más preferiblemente la cepa de cándida es una cepa *C. tropicalis* como se describió anteriormente. Se debe entender que una célula anfitriona que tiene un número de ocurrencia natural de los genes CYTb5 significa que abarca las células en donde ese número es cero (0).

En otro aspecto de la invención, se proporciona un método para incrementar la producción de un ácido dicarboxílico en una célula anfitriona. El método comprende (a) proporcionar la célula anfitriona que tiene ninguno, uno o más de los genes CYTb5; (b) incrementar, en la célula anfitriona, el número de genes CYTb5 que codifican una proteína CYTb5 que tiene una secuencia de aminoácido como se muestra en LAS Figuras 1A-1C en donde todos los anfitriones contienen una interrupción de la ruta de β -oxidación; y (c) cultivar la célula anfitriona en medio que contiene un sustrato orgánico que induce la expresión del gen CYTb5 para efectuar la producción incrementada de ácido dicarboxílico. Las células anfitrionas adecuadas incluyen aquellas como se describió anteriormente.

La célula anfitriona mencionada anteriormente utilizada para incrementar la producción de un ácido dicarboxílico también puede incluir un número incrementado de genes CPR o GYP comparados con la célula anfitriona tipo silvestre. Tales genes se han clonado y caracterizado como se describe en la patente U.S. No. 6,331,420 y WO 00/20566 A2 como se describió anteriormente. El cultivo de la célula anfitriona en medio que contiene un sustrato orgánico sobre regula el gen CPR, que resulta en producción incrementada de ácido dicarboxílico.

De acuerdo con lo anterior, la productividad (gramos de ácido dicarboxílico/litro/hr) de ácidos α , ω -dicarboxílicos sustancialmente puros, se incrementa significativamente al cultivar una célula anfitriona, por ejemplo, una cepa *C. tropicalis*, que tiene un mayor número de genes CPR y/o CYTb5 del tipo silvestre, y en los cuales los genes cromosómicos POX4A, POX4B y POX5 también se pueden interrumpir.

Un sustrato orgánico adecuado aquí puede ser cualquier compuesto orgánico que sea bio-oxidable con un ácido mono o poli-carboxílico. Tal un compuesto puede ser cualquier compuesto alifático saturado o insaturado o cualquier compuesto aromático carbocíclico o heterocíclico que tiene por lo menos un grupo metilo terminal, un grupo carboxilo terminal y/o un grupo funcional terminal que se puede oxidar con un grupo carboxilo por bio-oxidación. Un grupo funcional terminal que se deriva de un grupo carboxilo puede estar presente en la molécula de sustrato y se puede convertir a un grupo carboxilo mediante una reacción diferente de bio-oxidación. Por ejemplo, si el grupo terminal es un éster que ni el *C. tropicalis* tipo silvestre ni las modificaciones genéticas descritas aquí permitirían la hidrólisis de la funcionalidad éster a un grupo carboxilo, entonces se puede agregar una lipasa durante la etapa de fermentación para liberar los ácidos grasos. Los sustratos orgánicos adecuados incluyen, pero no se limitan a, ácidos grasos saturados, ácidos grasos insaturados, alcanos, alquenos, alquinos, y combinaciones de estos.

Los alcanos son un tipo de sustratos orgánicos saturados que son particularmente útiles aquí. Los alcanos pueden ser lineales o cíclicos, ramificados o de cadena recta, sustituidos o no sustituidos. Los alcanos particularmente preferidos son aquellos que tienen de aproximadamente 4 a aproximadamente 25 átomos de carbono, ejemplos de los cuales incluyen, pero no se limitan a, butano, hexano, octano, nonano, dodecano, tridecano, tetradecano, hexadecano, octadecano y similares.

Ejemplos de sustratos orgánicos insaturados que se pueden utilizar aquí incluyen, pero no se limitan a, olefinas internas tal como 2-penteno, 2-hexeno, 3-hexeno, 9-octadeceno y similares; ácidos carboxílicos insaturados tales como ácido 2-hexenoico y ésteres de estos, ácido oléico y ésteres de estos que incluyen ésteres de triglicerilo que tienen un contenido de ácido oléico relativamente alto, ácido erúsico y ésteres de este que incluyen ésteres de triglicerilo que tienen un contenido de ácido erúsico relativamente alto, ácido risinoléico y ésteres de este que incluyen ésteres de triglicerilo que tienen un contenido de ácido linoléico relativamente alto; alcoholes insaturados tal como 3-hexen-1-ol, 9-octadecen-1-ol y similares; Aldehídos insaturados tal como 3-hexen-1-al, 9-octadecen-1-al y similares. Adicionalmente, un sustrato orgánico que se puede utilizar aquí incluye compuestos alicíclicos que tienen por lo menos un doble enlace interno carbono-carbono y por lo menos un grupo terminal metilo, un grupo carboxilo terminal y/o un grupo funcional terminal que se puede oxidar con un grupo carboxilo mediante bio-oxidación. Ejemplos de

ES 2 298 276 T3

tales compuestos incluyen pero no se limitan a, 3,6-dimetilo, 1,4-ciclohexadieno, 3-metilciclohexeno, 3-metilo-1, 4-ciclohexadieno y similares.

Ejemplos de los compuestos aromáticos que se pueden utilizar aquí incluyen pero no se limitan a, arenos tal como o-, m-, p-xileno; o-, m-, ácido p-metil benzoico; piridinio dimetilo; esteroides y similares. El sustrato orgánico también puede contener otros grupos funcionales que son bio-oxidables con grupos carboxilo tal como un aldehído o un grupo alcohol. El sustrato orgánico también puede contener otros grupos funcionales que son bio-oxidables con grupos carboxilo y no interfieren con la bio-oxidación tal como halógenos, éteres, y similares.

Ejemplos de ácidos grasos saturados que se pueden aplicar a células que incorporan los genes CPR, CYP y/o CYTb5 de acuerdo con la presente invención incluyen ácidos caproico, enántico, caprílico, pelargónico, cáprico, undecílico, laurico, mirístico, pentadecanoico, palmítico, margarico, esteárico, araquídico, behénico y combinaciones de estos. Ejemplos de ácidos grasos insaturados que se pueden aplicar a las células que incorporan los presentes genes CPR y/o CYTb5 incluyen ácidos palmitoléicos, oléico, erúico, linoleico, linolénico y combinaciones de estos. Alcanos y fracciones de alcanos se pueden aplicar que incluyen enlaces de cadena de C12 a C24 en cualquier combinación. Un ejemplo de mezclas de ácidos grasos preferidos son Emersol®, Tallow y HOSFFA (aceite de girasol altamente oléico, es decir, mezcla de ácido graso que contiene aproximadamente 80% de ácido oléico disponible comercialmente como Edenor®) cada uno completamente disponible de Cognis Corp., Cincinnati, OH, la composición de ácido típicamente graso de Emersol® y Tallow es como sigue:

	TALLOW	EMERSOL
C14:0	3.5%	2.4%
C14:1	1.0%	0.7%
C15:0	0.5%	-----
C16:0	25.5%	4.6%
C16:1	4.0%	5.7%
C17:0	2.5%	-----
C17:1	-----	5.7%
C18:0	19.5%	1.0%
C18:1	41.0%	69.9%
C18:2	2.5%	8.8%
C18:3	-----	0.3%
C20:0	0.5%	-----
C20:1	-----	0.9%

Los siguientes ejemplos ilustran pero no limitan la invención. Todas las cepas microbianas relevantes se describen en la tabla 1 y 2, respectivamente. Tabla 1 lista de las Cepas *Escherichia coli* y *Candida tropicalis*.

ES 2 298 276 T3

CEPA E. Coli	GENOTIPO	FUENTE
XL1Blue-MRF'	endA1, gyrA96, hsdR17. lac ⁺ , recA1, relA1, supE44, thi-1, [F, lacI ^q Z M15, proAB, Tn 10]	Stratagene, La Jolla, CA
BM25.8	SupE44, thi (lac-proAB) [F' traD36, proAB ⁺ lacI ^q Z M15] l imm434 (kan ^R)P1 (cam ^R) hsdR (r _{k12} -m _{k12} -)	Clontech) Palo Alto, CA
XL0LR	(mcrA)183 (mcrCB-hsdSMR-mrr)173 endA1 thi-1 recA1 gyrA96 relA1 lac [F' proAB lacI ^q Z M15 Tn 10 (Tet ^r) Su ⁻ (no supresor I ^r (resistente a lambda)	Stratagene, La Jolla, CA
CEPA C. tropicalis	GENOTIPO	FUENTE
ATCC20336	Tipo silvestre	American Type Culture Collection, Rockville, MD
ATCC750	Tipo silvestre	American Type Culture Collection, Rockville, MD
ATCC 20962	ura3A / ura3B, pox4A::ura3A / pox4B::ura3A, pox5::ura3A / pox5::URA3A	Cognis
H5343 ura-	ura3A / ura3B, pox4A::ura3A / pox4B::ura3A, pox5::ura3A / pox5::URA3A, ura3-	Cognis
HDC11	ura3A / ura3B, pox4A::ura3A / pox4B::ura3A,	Cognis
	pox5::ura3A / pox5::URA3A,	Note- el HDC11-1 contiene menos copias integradas del inserto CYTb5 /CPR que el HDC11-2
	ura3::URA3A-CYTb5 + CPR B (CYTb5 y el CPR tienen orientación opuesta 5' a 3' con respecto uno del otro)	

TABLA 2

Lista de plásmidos aislados de librerías genómicas y construidas para uso en integraciones de gen

Plásmido	Vector base	Inserto	Tamaño del Inserto	tamaño del Plásmido	Descripción
pURAIin	pNEB193	URA3A	1706 bp	4399 bp	pNEB193 con el gen URA3A insertado en el sitio Asc I - Pme I, que genera un sitio Pac I
pURA REDB in	pURAIin	CPRB	3266 bp	7665 bp	pURAIin que contiene un alelo PCR CPRB que contiene los sitios de restricción Pac I
pHKM1	pTriplEx	Gen Truncado CPRA	Aprox. 3.8 kb	Aprox. 7.4 kb	Un gen CPRA truncado obtenido al seleccionar la primera librería que contiene la región 5' no traducida y el cuadro de lectura abierto 1.2 kb
pHKM4	pTriplEx	Gen Truncado CPRA	Aprox. 5 kb	Aprox. 8.6 kb	Un gen CPRA truncado obtenido al seleccionar la segunda librería que contiene la secuencia de extremo de la región 3' no traducida
pHKM9	pBC-CMV	Gen CPRB	Aprox. 5.3 kb	Aprox. 9.8 kb	Alelo CPRB aislado de la tercera librería

Ejemplo 1

Purificación de ADN genómico a partir de Candida tropicalis ATCC 20336

A. Construcción de librerías genómicas

50 ml de caldo de cultivo YEPD (ver apéndice) se inocula con una colonia única de *C. tropicalis* 20336 de Placa agar YEPD y se hace crecer durante la noche a 30°C. 5 ml del cultivo se inocula en 100 ml de caldo de cultivo YEPD fresco y se incuba a 30°C durante 4 a 5 hr con agitación. Las células se cosechan mediante centrifugación, se lavan dos veces con agua destilada estéril y se resuspenden en 4 ml de amortiguador de esferoplasto (1 M Sorbitol, 50 mM EDTA, 14 mM mercaptoetanol) se incuba durante 30 min a 37°C con agitación gentil. 0.5 ml de 2 mg/ml zimoliasa (ICN Pharmaceuticals, Inc., Irvine, CA) se agrega y se incuba a 37°C con agitación gentil durante 30 a 60 min. La formación de esferoplastos se monitorea mediante Lisis SDS. Los esferoplastos se cosechan por centrifugación breve (4,000 rpm, 3 min) y se lavan una vez con amortiguador de esferoplasto sin mercaptoetanol. Los esferoplastos cosechados se suspenden luego en 4 ml de amortiguador de lisis (0.2 M Tris/pH 8.0, 50 mM EDTA, 1% SDS) que contiene 100 mg/ml de RNasa (Qiagen Inc., Chatsworth, CA) se incuba a 37°C durante 30 a 60 min.

Las proteínas se desnaturalan y se extraen dos veces con un volumen igual de alcohol cloroformo/isoamilo (24:1) mediante agitación gentil de las dos fases por inversión manual. Las dos fases se separan por centrifugación a 10,000 rpm durante 10 min y la fase acuosa que contiene el ADN de alto peso molecular se recupera. Para la capa acuosa de NaCl se agrega una concentración final de 0.2 M y el ADN se precipita al agregar 2 vol de etanol. El ADN precipitado se agita con una vara de vidrio limpia y se resuspende en amortiguador TE r (1Q mM Tris/pH 8.0, 1 mM EDTA) y se le permite disolver durante la noche a 4°C. Al ADN disuelto, la RNasa libre de cualquier actividad DNasa (Qiagen Inc., Chatsworth, CA) se agrega a una concentración final de 50 mg/ml se incuba a 37°C durante 30 min. Luego la proteasa (Qiagen Inc., Chatsworth, CA) se agrega a una concentración final de 100 mg/ml se incuba a 55 a 60°C durante 30 min. La solución se extrae una vez con un volumen igual de alcohol fenol/cloroformo/isoamilo (25:24:1) y una vez con volumen igual de alcohol de cloroformo/isoamilo (24:1). A la fase acuosa 0.1 vol de 3 M acetato de sodio y 2 volúmenes de etanol enfriado con hielo (prueba 200) se agregan y el ADN de alto peso molecular se agita con una vara de vidrio se disuelve en 1 a 2 ml de amortiguador TE.

ES 2 298 276 T3

B. Preparación de ADN genómico para PCR

Amplificación de los genes *CYTb5* y *CPR*

- 5 Cinco 5 ml de Medio YPD se inocula con una colonia única y se hace crecer a 30°C durante la noche. El cultivo se centrifuga durante minutos a 1200 x g. El super nadante se remueve por aspiración y 0.5 ml de una solución de sorbitol (0.9 M sorbitol, 0.1 M Tris-Cl pH 8.0, 0.1 M EDTA) se agrega al gránulo. El gránulo se resuspende al centrifugar y se agrega 1 ml de 2-mercaptoetanol y 50 ml de solución de zimoliasa 10 mg/ml a la mezcla. El tubo se incuba a 37°C durante 1 hr en un agitador rotatorio (200 rpm). El tubo se centrifuga luego durante 5 min a 1200 x g y el super nadante se remueve por aspiración. El gránulo del protoplasto se resuspende en 0.5 ml 1x TE (10 mM Tris-Cl pH 8.0, 1 mM EDTA) y se transfiere a un tubo de microcentrifuga de 1.5 ml. Los protoplastos se lisan mediante la adición de 50 ml. 10% SDS seguido por incubación a 65°C durante 20 min. Luego, 200 ml de 5M de acetato de potasio se agregan y después de mezcla, el tubo se incuba en hielo durante por lo menos 30 min. Los residuos celulares se remueven mediante centrifugación a 13,000 x g durante 5 min. El super nadante se remueve cuidadosamente y se transfiere un nuevo tubo microfugo. El ADN se precipita por la adición de 1 ml 100% (200 prueba) de etanol seguido por centrifugación durante 5 min a 13,000 x g. El gránulo de ADN se lava con 1 ml 70% etanol seguido por centrifugación durante 5 min a 13,000 x g. Después de secar parcialmente el ADN bajo vacío, se resuspende en 200 ml de 1x TE. La concentración de ADN se determina mediante la proporción de la absorbancia en 260 nm/280 nm ($A_{260/280}$).

Ejemplo 2

Construcción de librerías genómicas *Candida tropicalis* 20336

- 25 Se construyen Tres librerías genómicas de *C. tropicalis*, dos en Clontech Laboratories, Inc., (Palo Alto, CA) y una en Henkel Corporation (Cincinnati, OH).

A. Librerías Clontech

- 30 La primer librería Clontech se hace como sigue: ADN genómico, se prepara de *C. tropicalis* 20336 como se describió anteriormente, se digiere parcialmente con EcoRI y se fracciona en partes por electroforesis de gel para eliminar los fragmentos más pequeños de 0.6 kb. Luego del fraccionamiento del tamaño, varias ligaciones de los fragmentos de ADN genómico Eco RI y lambda (λ) brazos vector TriplEx™ (Figura 16) con extremos adhesivos Eco RI que se empa- can en las cabezas de fago λ bajo condiciones diseñadas para obtener un millón de clones independientes. La segunda librería genómica se construye como sigue: el ADN genómico se digiere parcialmente con Sau3A1 y se fracciona en partes mediante electroforesis de gel. Los fragmentos de ADN tienen extremos rombos utilizando protocolos estándar como se describe, por ejemplo, en Sambrook *et al*, Molecular Cloning: A Laboratory Manual, 2ed. Cold Spring Harbor Press, USA (1989). La estrategia es llenar los restos del Sau3A1 con polimerasa Klenow (Life Technologies, Grand Island, NY) seguido por digestión con nucleasa S1 (Life Technologies, Grand Island, NY). Después de la digestión de nucleasa S1 los fragmentos se llenan una o más veces con polimerasa Klenow para obtener los fragmentos de ADN de extremo romo. Los ligadores EcoRI se ligan con estos fragmentos de ADN de extremo romo seguido por ligación en el vector λ TriplEx. La librería resultante contiene aproximadamente 2×10^6 clones independientes con un tamaño de inserto promedio de 4.5 kb.

45 B. Librería Cognis

- Se construye una librería genómica utilizando el vector λ ZAP Express™ (Stratagene, La Jolla, CA) (Figura 17A). El ADN genómico se digiere parcialmente con Sau3A1 y fragmentos en el rango de 6 a 12 kb se purifican de un gel agarosa después de electroforesis del ADN digerido. Estos fragmentos de ADN se ligan luego con brazos de vector λ ZAP Express™ digerido Bam HI de acuerdo con los protocolos del fabricante. Las tres ligaciones se establecen para obtener aproximadamente 9.8×10^5 clones independientes. Las tres librerías se agrupan y amplifican de acuerdo con las instrucciones del fabricante para obtener un depósito de alto título ($>10^9$ unidades que forman placa/ml) durante almacenamiento a largo plazo. El título de la librería fago empackado se comprobó después de la infección de células *E. coli* XL1Blue-MRF. El *E. coli* XL1Blue-MRF' se hace crecer durante la noche en medio LB o, NZCYM (Apéndice) que contiene 10 mM $MgSO_4$ y 0.2% de maltosa a 37°C o 30°C, respectivamente con agitación. Las células se centrifugan luego y se resuspenden en 0.5 a 1 volumen de 10 mM de $MgSO_4$. 200 ml de este cultivo *E. coli* se mezclan con varias diluciones de librería fago empackadas y se incuban a 37°C durante 15 min. A esta mezcla 2.5 ml de azarosa superior LB o azarosa superior NZCYM (mantenida 60°C) (ver apéndice) se agregan y la placa en agar LB o agar NCZYM (ver apéndice) presente en platos petri de 82 mm. El fago se le permite propagarse durante la noche a 37°C para obtener placas discretas y se determina el título fago.

Ejemplo 3

65 Selección de librerías genómicas

Los vectores λ TriplEx™ y λ ZAP Express™ son vectores fagémidos que se pueden propagar como fago o plásmidos de ADN (después de conversión de fago a plásmidos). Por lo tanto, las librerías genómicas construidas en estos

vectores se pueden seleccionar mediante hibridación por placa (selección de la forma lambda de la librería) o por hibridación de colonia (selección de la forma de plásmido de la librería después de conversión de fago a plásmido). Ambos vectores son capaces de expresar los genes clonados y la principal diferencia es el mecanismo de división del plásmido del fago de ADN. El sitio de clonación en λ TriplExTM se ubica dentro de un plásmido que está presente en el fago y se flanquea por el sitio loxP (Figura 16). Cuando el λ TriplExTM se introduce en la cepa *E. coli* BM25.8 (suministrada por Clontech), la recombinasa Cre presente en promotores BM25.8 la división y la circulación de plásmido pTriplEx del fago λ TriplExTM a en los sitios loxP. El mecanismo de división del plásmido pBK-CMV (Figura 17B) del fago λ ZAP ExpressTM es diferente. Se requiere la ayuda de un fago auxiliar de tal manera que ExAssistTM (Stratagene) y una cepa *E. coli* tal como XLOR (Stratagene). pTriplEx y pBK-CMV pueden replicar autónomamente en *E. coli*.

A. Selección de librerías genómicas (forma de Plásmido)

1) levantamiento de colonias

Una colonia única de *E. coli* BM25.8 se inocula en 5 ml de LB que contiene 50 mg/ml de canamicina, 10 mM MgSO₄ y 0.1% de maltosa y se hace crecer durante la noche a 31°C, 250 rpm. A 200 ml de este cultivo durante la noche e (~4 X 10⁸ células) 1 ml de librería fago (2 - 5 X 10⁶ unidades formadoras de placa) y 150 ml de caldo de cultivo LB se agregan y se incuban a 31°C durante 30 min después de lo cual 400 ml de Caldo de cultivo LB se agrega y se incuba a 31°C, 225 rpm durante 1 h. Este cultivo bacteriano se diluye y se coloca en placas LB de agar que contienen 50 mg/ml de ampicilina (Sigma Chemical Company, St. Louis, MO) y canamicina (Sigma Chemical Company) para obtener 500 a 600 colonias/placa. Las placas se incuban a 37°C durante 6 a 7 horas hasta que las colonias llegan a ser visibles. Las placas se almacenan luego a 4°C durante 1.5 hrs antes de colocar un disco de membrana de transferencia de hibridación Colony/Plaque ScreenTM (DuPont NEN Research Products, Boston, MA) sobre la placa en contacto con las colonias bacterianas. La transferencia de colonias a la membrana se permite para proceder durante 3 a 5 min. La membrana se levanta luego y se coloca en una placa LB fresco de agar (ver apéndice) que contiene 200 mg/ml de cloramfenicol con el lado expuesto a las colonias bacterianas que lo enfrentan. Las placas que contienen las membranas se incuban luego a 37°C durante la noche con el fin de permitir el desarrollo completo de las colonias bacterianas. Las placas de agar LB de las cuales las colonias se levantan inicialmente se incuban a 37°C durante la noche y se almacenan a 4°C para uso futuro. A la mañana siguiente las membranas que contienen las colonias bacterianas se levantan y se colocan en dos hojas de papel Whatman 3M (Whatman, Hillsboro, OR) saturado con 0.5 N NaOH y se dejan a temperatura ambiente (RT) durante 3 a 6 min para lisar las células. El tratamiento adicional de las membranas se describe en el protocolo proporcionado por NEN Research Products.

2) Hibridaciones de ADN

Se secan membranas durante la noche antes de hibridar para sondas de oligonucleótido preparadas utilizando un sistema de detección ECLTM 3'-oligoetiquetado no radioactivo de Amersham Life Sciences (Arlington Heights, IL). El etiquetado de ADN, la pre-hibridación y las hibridaciones se desarrollan de acuerdo con los protocolos del fabricante. Después de hibridación, las membranas se lavan dos veces a temperatura ambiente en 5 X SSC, 0.1% SDS (en un volumen equivalente a 2 ml/cm² de membrana) durante 5 min cada uno seguido por dos lavadas a 50°C en 1X SSC, 0.1% SDS (en un volumen equivalente a 2 ml/cm² de membrana) durante 15 min cada uno. La señal de hibridación se genera y se detecta luego con el Hyperfilm ECLTM (Amersham) de acuerdo con los protocolos del fabricante. Las membranas se alinean con placas que contienen las colonias bacterianas de las cuales las colonias que se levantan son desarrolladas y las colonias que corresponden a señales positivas en rayos X se aíslan luego y cultivan en caldo de cultivo LB. Los plásmidos de ADN se aíslan de estos cultivos y se analizan para restricción de digestión de enzima por secuenciamiento de ADN.

B. Selección de librerías genómicas (forma placa)

1) Forma de placa de librería λ

Se hacen crecer células de *E. coli* XL1Blue-MRF durante la noche en medio LB (25 ml) que contiene 10 mM MgSO₄ y 0.2% de maltosa a 37°C, 250 rpm. Las células se centrifugan luego (2,200 x g durante 10 min) y se resuspende en 0.5 volúmenes de 10 mM MgSO₄. 500 ml de este cultivo *E. coli* se mezcla con una suspensión de fago que contiene 25,000 partículas de fago lambda amplificadas y se incuban a 37°C durante 15 min. A esta mezcla 6.5 ml de agarosa superior NZCYM (mantenida a 60°C) (ver apéndice) se agrega y se coloca en placas en 80 - 100 ml de NCZYM de agar (ver apéndice) presente en un plato petri de 150 mm. Al fago se le permite propagar durante la noche a 37°C para obtener placas discretas. Después que se hacen crecer durante la noche se almacenan las placas en un refrigerador durante 1-2 horas antes que se desarrolle el levantamiento de placa.

2) Levantamiento de placa e hibridaciones de ADN

Se colocan membranas de nylon Magna LiftTM (Micron Separations, Inc., Westborough, MA) sobre la superficie de agar en contacto completo con una placa y se permite la transferencia de placas a membranas de nylon para proceder durante 5 min a RT. Después de la transferencia de placa la membrana se coloca en dos láminas de papel de filtro de Whatman 3MTM (Whatman, Hillsboro, OR) saturado con 0.5 N de NaOH, 1.0 M de solución NaCl y se dejan durante 10 min a RT para desnaturalizar ADN. El exceso de solución desnaturalizante se remueve brevemente por transferencia en papel Whatman 3M. Las membranas se transfieren luego a dos hojas de papel Whatman 3MTM saturadas con 0.5 M

ES 2 298 276 T3

Tris-HCl (pH 8.0), 1.5 M de NaCl y se dejan durante 5 min para neutralizar. Las membranas se lavan brevemente en 200 - 500 ml de 2 X SSC, se secan por aire y se dejan reposar durante 30 - 40 min a 80°C. Las membranas se sondan luego con ADN etiquetado.

- 5 Las membranas se pre-lavan con una solución de 200 - 500 ml de 5 X SSC, 0.5% SDS, 1 mM EDTA (pH 8.0) durante 1 - 2 hr a 42°C con agitación (60 rpm) para librarse de los residuos bacterianos de las membranas. Las membranas se pre-hibridan durante 1 - 2 hrs a 42°C con (en un volumen equivalente a 0.125 - 0.25 ml/cm² de membrana) amortiguador ECL Gold™ (Amersham) que contiene 0.5 M NaCl y 5% de reactivo bloqueador. Los fragmentos de ADN que se utilizan como sondas se purifican del gel azarosa utilizando un kit de extracción de gel
- 10 QIAEX II™ (Qiagen Inc., Chatsworth, CA) de acuerdo con el protocolo del fabricante y etiquetado utilizando el kit de etiquetado de ácido nucleico directo Amersham ECL™ (Amersham). El etiquetado, ADN (solución de hibridación 5-10 ng/ml) se agrega a las membranas pre-hibridadas y la hibridación se permite proceder durante la noche. Al siguiente día se lavan las membranas con agitación (60 rpm) dos veces a 42°C durante 20 min cada vez en (en un volumen equivalente a 2, ml/cm² de membrana) un amortiguador que contiene 0.1 (alta restricción) o 0.5 (baja restricción) X
- 15 SSC, 0.4% SDS y 360 g/l de urea. Esto es seguido por dos lavados de 5 min a temperatura ambiente en (en un volumen equivalente a 2 ml/cm² de membrana) 2 X SSC. Se generan señales de hibridación utilizando reactivo de detección de ácido nucleico ECL™ y se detecta utilizando Hyperfilm ECL™ (Amersham).

- 20 Tapones de Agar que contienen placas que corresponden a señales positivas en la película de rayos X se toman a partir de placas maestras utilizando el extremo amplio de la pipeta Pasteur. Las placas se seleccionan al alinear las placas con la película de rayos X. En esta etapa, se toman generalmente múltiples placas. Las partículas fago se eluyen a partir de tapones de agar al enjuagar en 1 ml de amortiguador SM (Sambrook *et al.*, *supra*) durante la noche. El eluido fago se diluye luego y se colocan placas con células *E. coli* XL1Blue-MRF que crecen frescas para obtener 100 - 500 placas por 85 mm de placa agar NCZYM. Las placas se transfieren a membranas de nylon Magna Lift como se hizo
- 25 anteriormente y se sondan de nuevo utilizando la misma sonda. Las placas sencillas se aíslan bien correspondiendo a señales en películas de rayos X que se recogen al remover los tapones de agar y eluir el fago mediante enjuague durante la noche en 0.5 ml de amortiguador SM.

C. Conversión de clones λ a la forma de plásmido

- 30 Clones lambda aislados se convierten a la forma de plásmido para análisis adicional. La conversión de la placa a la forma de plásmido se logra al infectar las placas con la cepa *E. coli* strain BM25.8. La cepa *E. coli* se hace crecer durante la noche a 31°C, 250 rpm en Caldo de cultivo LB que contiene 10 mM MgSO₄, 0.2% de maltosa hasta que se alcanza OD₆₀₀ de 1.1-1.4. Se remueven diez mililitros de cultivo durante la noche y se mezcla con 140 ml de 1 M de MgCl₂. Un volumen de 200 ml de células se remueve, se mezcla con 150 ml de suspensión de fago e incuban a
- 35 31°C durante 30 min. El Caldo de cultivo LB (400 ml) se agrega al tubo y se continúa la incubación a 31°C durante 1 hr con agitación, 250 rpm. 1 - 10 ml de la suspensión celular infectada se colocan placas en agar LB que contiene 100 mg/ml ampicilina (Sigma Chemical Company, St. Louis, MO). Las colonias bien aisladas se recogen y se hacen crecer durante la noche en 5 ml de Caldo de cultivo LB que contiene 100 mg/ml ampicilina a 37°C, 250 rpm. El ADN de Plásmido se aísla de estos cultivos y se analiza. Para convertir el vector λ ZAP Express™ a la forma de plásmido se utilizan las cepas *E. coli* XL1Blue-MRF⁺ y XLOR. La conversión se desarrolla de acuerdo con los protocolos del fabricante (Stratagene) para división de placa única.
- 40

45 Ejemplo 4

Transformación de *C. tropicalis* utilizando acetato de litio

- Se utiliza el siguiente protocolo para transformar *C. tropicalis* de acuerdo con el procedimiento descrito en Current
- 50 Protocols in Molecular Biology, Supplement 5, 13.7.1 (1989).

- 5 ml de YEPD se inocula con *C. tropicalis* H5343 ura de materia prima congelada y se incuba durante la noche en un agitador New Brunswick a 30°C y 170 rpm. Al siguiente día, 10 μ l del cultivo se inocula en 50 ml YEPD y se hace crecer y se continúa a 30°C, 170 rpm. Al siguiente día las células se cosechan en un OD₆₀₀ de 1.0. El cultivo se transfiere un tubo de polipropileno de 50 ml y se centrifuga a 1000 X g durante 10 min. El gránulo celular se resuspende en 10
- 55 ml estéril TE (10 mM Tris-Cl and 1 mM EDTA, pH 8.0). Las células se centrifugan de nuevo en 1000 X g durante 10 min y el gránulo celular se resuspende en 10 ml de una solución de acetato de litio estéril [LiAc (0.1 M de acetato de litio, 10 mM Tris-Cl, pH 8.0, 1 mM EDTA)]. Luego de centrifugación a 1000 X g durante 10 min., el gránulo se resuspende en 0.5 ml LiAc. Esta solución se incuba durante 1 hr a 30°C mientras se agita gentilmente a 50 rpm. Una
- 60 alícuota de 0.1 ml de esta suspensión se incuba con 5 μ g de ADN transformante a 30°C sin agitación durante 30 min. 0.7 ml de solución PEG (40% p/v de polietilenglicol 3340, 0.1 M de acetato de litio, 10 mM de Tris-Cl, pH 8.0, 1 mM EDTA) se agrega se incuba a 30°C durante 45 min. Los tubos luego se colocan a 42°C durante 5 min. Una alícuota de 0.2 ml se coloca en medio completo sintético menos uracilo (SC - uracilo) (Kaiser *et al.* Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press, USA. 1994). El crecimiento de los transformantes se monitorea durante 5 días.
- 65 Después de tres días, varios transformantes se recogen y se transfieren a placas SC-uracilo para preparación de ADN genómico y selección.

ES 2 298 276 T3

Ejemplo 5

Aislamiento de ADN de Plásmido

- 5 El ADN del Plásmido se aísla de cultivos de *E. coli* utilizando el kit de aislamiento de plásmido Qiagen (Qiagen Inc., Chatsworth, CA) de acuerdo con las instrucciones del fabricante.

Ejemplo 6

10

Secuenciamiento de ADN y análisis

- Se desarrolla el secuenciamiento de ADN en Sequetech Corporation (Mountain View, CA) utilizando el secuenciador automatizado Applied Biosystems (Perkin Elmer, Foster City, CA). Las secuencias de ADN se analizan con los paquetes de software MacVector y Gene Works (Oxford Molecular Group, Campbell, CA).

Ejemplo 7

Protocolos PCR

- Se lleva a cabo amplificación PCR en un termociclizador Perkin Elmer utilizando el kit de enzima Ampli Taq Gold (Perkin Elmer Cetus, Foster City, CA) de acuerdo con las especificaciones del fabricante. La siguiente amplificación exitosa, en algunos casos, los productos se digieren con las enzimas apropiadas y se purifican con gel utilizando QiaexII (Qiagen, Chatsworth, CA) según las instrucciones del fabricante. En casos específicos la polimerasa Ultima Taq (Perkin Elmer Cetus, Foster City, CA) o la polimerasa Expand Hi-Fi Taq (Boehringer Mannheim, Indianapolis, IN) se utilizan según las recomendaciones del fabricante o como se define en la tabla 3.

30

TABLA 3

Condiciones de amplificación PCR utilizadas con diferentes combinaciones de cebador

35	COMBINACIÓN DE CEBADOR	Taq	CONDICIÓN DE DESNATURALIZACIÓN DE PLANTILLA	TEMPERTURA/TIEMPO DE HIBRIDACIÓN	TEMPERATURA/TIEMPO DE EXTENSIÓN	CICLO número
40	3674-41-1/ 41-2/ 41-4 +3674-41-4	Ampli-Taq Gold	94 C/30 seg.	55 C/30 seg.	72 C/1 min	30
45	URA cebador 1a	Ampli-Taq Gold	95 C/1 min	70 C/1 min	72 C/2 min	35
50	URA cebador 1b	Ampli-Taq Gold	95 C/1 min	70 C/1 min	72 C/2 min	35
55	URA cebador 2a	Ampli-Taq Gold	95 C/1 min	70 C/1 min	72 C/2 min	35
60	URA cebador 2b	Ampli-Taq Gold	95 C/1 min	70 C/1 min	72 C/2 min	35
	CPR B#1	Expand	94 C/15 seg.	50 C/30 seg.	68 C/3 min	10
	CPR B#2	Hi-Fi Taq	94 C/15 seg.	50 C/30 seg.	68 C/3 min +20 seg/ciclo	15
	CYTb5 #1	Ultma	95°C/15 seg.	45°C/30 seg.	72°C/2 min	25
	CYTb5 #2	Taq				

- La Tabla 4 más adelante contiene una lista de cebadores utilizados para la amplificación PCR para construir vectores de integración de gen o para generar sondas para detección y aislamiento de gen.

65

Tabla 4. Tabla de cebador para amplificación PCR para construcción de vectores de integración de gen, para generar sondas, para detección y aislamiento de gen y para obtener secuencias de ADN de construcciones. (A-deoxiadenosino trifosfato [dATP], G-deoxiguanosina trifosfato [dGTP], C-deoxicitosina trifosfato [dCTP], T-deoxitimidina trifosfato [dTTP], Y-dCTP o dTTP, R-dATP o dGTP, W-dATP o dTTP, M-dATP o dCTP).

Genes objetivo	Nombre cebador patente	Nombre del cebador Lab	Secuencia (5'a 3')	Tamaño producto PCR
CPRB	CPRB#1	3698-20A	CCTTAA TTAAGAGGTCGTTGGTTGAGTTTTC (SEQ. ID NO. 17)	3266 bp
	CPRB#2	3698-20B	CCTTAA TTAATTGATAATGACGTTGCGGG (SEQ. ID NO. 18)	
URA3A	URA Primer 1a	3698-7C	AGCGCGCGGAGTCCAAAAGACCAACCTCTG (SEQ. ID NO. 19)	956 bp
	URA Primer 1b	3698-7D	CCTTAA TTAATACGTGGATACCTTCAAGCAAGTG (SEQ. ID NO. 20)	
URA3A	URA Primer 2a	3698-7A	CCTTAA TTAAGCTCAGGATTTTGGGATTTTCGAG (SEQ. ID NO. 21)	750 bp
	URA Primer 2b	3698-7B	GGGTTTAAACCGCAGAGGTTGGTCTTTTGGACTC (SEQ. ID NO. 22)	
			GGGTTTAAAC - Pme I sitio de restricción (SEQ. ID NO. 23)	
			AGCGCGCGCC - Asc I sitio de restricción (SEQ. ID NO. 24)	
			CCTTAA TTAAC - Pac I sitio de restricción (SEQ. ID NO. 25)	
CPR	FMN1	3674-41-1	TCYCAACWGGTACWOCWGAA (SEQ. ID NO. 26)	
CPR	FMN2	3674-41-2	GGTTGGGTAATCWACTTAT (SEQ. ID NO. 27)	
CPR	FAD	3674-41-3	CGTTATTATTCYATTCTTC (SEQ. ID NO. 28)	
CPR	NADPH	3674-41-4	GCMACACCGGTACCTGGACC (SEQ. ID NO. 29)	

continuación

Genes objetivo	Nombre cebador patente	Nombre del cebador Lab	Secuencia (5'a 3')	Tamaño producto PCR
CPR	PRK1.F3	PRK1.F3	ATCCCAATCGTAATCAGC (SEQ. ID NO. 30)	
CPR	PRK1.F5	PRK1.F5	ACTGTCTTCGTTTAGCA (SEQ. ID NO. 31)	
CPR	PRK4.R20	PRK4.R 20	CTACGTCGTGTGTGATGC (SEQ. ID NO. 32)	
pTriplEx vector	Triplex5'	Triplex5'	CTCGGGAAGCGCGCCATTGTGTGG (SEQ. ID NO. 33)	
pTriplEx vector	Triplex3'	Triplex3'	TAATACGACTCACTATAGGGCGAATTGGC (SEQ. ID NO. 34)	
CYTb5	CYTb5 #1		GGGTTAATTAAACATACTTCAAGCAGTTTGG (SEQ. ID NO. 35)	
	CYTb5 #2		CCCTTAATTAAGGGGGGATGGAAGTGCGCG (SEQ. ID NO. 36)	
	3698-66A		ATAAGATGCGGCCGCTGAACGAGAACCACATCCAGGAG (SEQ. ID NO. 37)	
	3698-66B		CCTTAATTAAGGATAACCAGATCCATACGTGCG (SEQ. ID NO. 38)	

Ejemplo 8

Procedimiento PCR para colonia de levadura para confirmación de integración de gen en el genoma de C. tropicalis

- 5 Las colonias de levadura únicas se remueven de la superficie de placas de transformación, se suspenden en 50 μ l de amortiguador de esferoplasto (50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.0 mg/ml Zimoliasa, 5% de glicerol) se incuban a 37°C durante 30 min. Luego de incubación, la solución se calienta durante 10 min a 95°C para lisar las células. Cinco μ l de esta solución se utiliza como una plantilla en PCR. La polimerasa Hi-Fi Taq expandida (Boehringer Mannheim, Indianapolis, IN) se utiliza en PCR acoplada con un cebador específico de gen (gen a ser integrado) y un cebador URA3. Si no ocurre la integración, la amplificación produciría un producto PCR de tamaño predicho que confirma la presencia de un gen integrado.

Ejemplo 9

- 15 *Métodos de fermentación para estudios de inducción de gen*

- Se carga un fermentador con un medio de crecimiento semi sintético que tiene la composición 75 g/l de glucosa (anhidra), 6.7 g/l de base de nitrógeno de levadura (Difco Laboratories), 3 g/l de extracto de levadura, 3 g/l de sulfato de amonio sulfato, 2 g/l de fosfato monopotásico, 0.5 g/l de cloruro de sodio. Los Componentes se hacen como soluciones concentradas para autoclave luego de agregar el fermentador después de enfriamiento: pH final aproximadamente 5.2. Esta carga se inocula con 5-10% de un cultivo durante la noche de *C. tropicalis* ATCC 20962 preparado en medio YM (Difco Laboratories) como se describe en los métodos de los ejemplos 17 y 20 de la patente U.S. No. 5,254,466. El *C. tropicalis* ATCC 20962 es un POX 4 y POX 5 interrumpido de *C. tropicalis* ATCC 20336. Se suspende el aire y la agitación suministrada para mantener y disolver el oxígeno en más de aproximadamente 40% de saturación versus aire. El pH se mantiene en aproximadamente 5.0 a 8.5 por la adición de 5N soda cáustica en pH de control. una corriente de alimentación de ácido graso (ácido oleico comercial en este ejemplo) que tiene una composición típica: 2.4% C₁₄; 0.7% C_{14:1}; 4.6% C₁₆; 5.7% C_{16:1}; 5.7% C_{17:1}; 1.0% C₁₈; 69.9% C_{18:1}; 8.8% C_{18:2}; 0.30% C_{18:3}; 0.90% C_{20:1} y una carga de co-sustrato de glucosa se agregan en un modo de tanda de alimentación que comienza cerca del final del crecimiento exponencial. El cáustico se agrega sobre control de pH durante la bio-conversión de ácidos grasos a diácidos para mantener el pH en el rango deseado. Típicamente, las muestras para estudios de inducción de gen se recolectan justo antes del inicio de la admisión de ácidos grasos y durante las primeras diez horas de bio-conversión. La determinación de ácido graso y contenido de diácido se determina mediante un protocolo de metil éster estándar utilizando cromatografía líquida de gas (GLC) (ver Ejemplo 12). Se mide la inducción de gen utilizando el protocolo QC-RT-PCR descrito en esta solicitud.

Ejemplo 10

- 40 *Preparación de ARN*

- La primera etapa de este protocolo involucra el aislamiento del ARN celular total de cultivos de *C. tropicalis*. El ARN celular se aísla utilizando el mini kit Qiagen RNeasy (Qiagen Inc., Chatsworkh, CA) como sigue: se recolectan muestras de 2 ml de *C. tropicalis* del fermentador en tubos de estilo Eppendorf de tapa rosca de 2 ml estándar en varios momentos antes y después de la adición del ácido graso o sustrato alciano. Las muestras celulares se congelan inmediatamente en nitrógeno líquido o un baño de hielo/alcohol después de su cosecha del fermentador. Para aislar el ARN total de las muestras, se le permite a los tubos congelarse en hielo y las células granuladas por centrifugación en una micro fuga durante 5 min a 4°C y el super nadante se descarta mientras se mantiene el gránulo frío en hielo. Los tubos de microfuga se llenan 2/3, con glóbulos de Zirconia/Sílice enfriados con hielo (0.5 nun de diámetro, Biospec Products, Bartlesville, OK) y el tubo se llena en la parte superior con amortiguador de lisis RLT* enfriado con hielo (* amortiguador incluido con el mini kit Qiagen RNeasy). La ruptura celular se logra al colocar las muestras en un mezclador de mini glóbulo (Biospec Products, Bartlesville, OK) e inmediatamente homogenizado a velocidad completa durante 2.5 min. Las muestras se les permite enfriar en un baño de agua con hielo durante 1 min y el proceso de homogenización/enfriamiento se repite dos veces más durante un total de 7.5 min de homogenización en un mezclador de glóbulos. Las muestras celulares homogenizadas se microfugan a velocidad completa durante 10 min y 700 ml del super nadante que contiene ARN removido y transferido a un nuevo tubo eppendorf. Se agrega 700 ml de 70% etanol a cada muestra seguido por mezcla por inversión. Esta y todas las etapas posteriores se desarrollan a temperatura ambiente. Setecientos microlitros de cada muestra tratada con etanol se transfieren a una columna de spin Qiagen RNeasy, seguido por centrifugación a 8,000 x g durante 15 seg. El flujo se descarta y la columna se recarga con la muestra restante (700 ml) y se re-centrifuga a 8,000 x g durante 15 seg. La columna se lava una vez con 700 ml de amortiguador RW1*, y se centrifuga a 8,000 x g durante 15 seg. Y se descarta el flujo. La columna se coloca en un nuevo tubo de recolección de 2 ml y se lava con 500 ml de amortiguador RPE* y se descarta el flujo. El lavado RPE* se repite con centrifugación en 8,000 x g durante 2 min y se descarta el flujo. La columna spin se transfiere a un nuevo tubo de recolección de 1.5 ml y 100 ml de RNasa libre de agua se agrega a la columna seguido por centrifugación a 8,000 x g durante 15 segundos. Se agrega un adicional de 75 ml de agua libre de RNasa a la columna seguido por centrifugación en 8,000 x g durante 2 min. El ARN eluido en el flujo de agua se recolecta para purificación adicional.

El eluado de ARN se trata luego para remover el ADN contaminante. Veinte microlitros de amortiguador de 10X DNasa I (0.5 M tris (pH 7.5), 50 mM CaCl₂, 100 mM MgCl₂), 10 ml de DNasa libre de RNasa 1 (2 Unidades/ml, Ambion Inc., Austin, Texas) y 40 unidades de Rnasin (Promega Corporation, Madison, Wisconsin) se agregan a la muestra de ARN. La mezcla se incuba luego a 37°C durante 15 a 30 min. Las muestras se colocan en hielo y se agrega 250 µl de Amortiguador de lisis RLT* y 250 µl etanol (200 prueba). Las muestras se mezclan luego por inversión. Las muestras se transfieren a columnas spin Qiagen RNeasy y se centrifugan a 8,000 x g durante 15 seg. Y se descarta el flujo. Las columnas se colocan en tubos de recolección nuevos de 2 ml y se lavan dos veces con 500 µl de amortiguador de lavado RPE* y se descarta el flujo. Las columnas se transfieren a tubos de 1.5 ml eppendorf nuevos y se elue el ARN mediante la adición de 100 µl de agua tratada con DEPC seguido por centrifugación a 8,000 x g durante 15 seg. El ARN Residual se recolecta al agregar 50 µl adicional de agua libre de RNasa a la columna spin seguido por centrifugación a velocidad completa durante 2 min. 10 µl de la prelación de ARN se remueve y se cuantifica mediante el método (A_{260/280}). El ARN se almacena a - 70°C. Se encuentra que los rendimientos son de 30-100 mg de ARN por 2.0 ml de caldo de cultivo de fermentación.

Ejemplo 11

Protocolo de reacción de cadena de polimerasa de transcripción inversa competitiva cuantitativa (QC-RT-PCR)

El QC-RT-PCR es una técnica utilizada para cuantificar la cantidad de una ARN específico en una muestra de ARN. Esta técnica emplea la síntesis de una molécula de ADN específica que es complementaria con una molécula de ARN en la muestra original mediante transcripción inversa y su posterior amplificación por reacción de cadena de polimerasa. Mediante la adición de varias cantidades de una molécula de ARN competidor con la muestra uno puede determinar la concentración de la molécula de ARN de interés (por ejemplo, la transcripción mARN del gen CYTb5). Los niveles de transcripciones de mARN específicos se ensayan durante el tiempo en respuesta a la adición de ácido graso o sustratos alceno para el medio de crecimiento de cultivos *C. tropicalis* de crecimiento de fermentación para la identificación y caracterización de los genes involucrados en la oxidación de estos sustratos. Este enfoque se utiliza para identificar el gen CYTb5 involucrado en la oxidación de cualquier sustrato dado con base en su regulación transcripcional.

A. Diseño de cebador

El primer requerimiento para el QC-RT-PCR es el diseño de los pares cebador a ser utilizados en la transcripción inversa y la posterior reacción PCR. Estos cebadores necesitan ser únicos y específicos para el gen de interés. Los cebadores utilizados para medir la expresión del gen CYTb5 de *C. tropicalis* 20336 que utilizan el protocolo QC-RT-PCR se listan en la tabla 5.

TABLA 5

Cebadores utilizados para medir la expresión del gen *C. tropicalis* CYTB5 en las reacciones QC-RT-PCR

Nombre del cebador	Dirección	Objetivo	Secuencia
3740-179A (SEQ. ID NO. 39)	F	CYTb5	CACACCACCCACGACGACTTGTG
3740-179C (SEQ. ID NO. 40)	B	CYTb5	CTTCCGTGCTGAACGACTGCG
F=Delantero B=Trasero			

B. Diseño y síntesis de la plantilla de ADN competidor

El ARN competidor se sintetiza *in Vitro* a partir de una plantilla de ADN competidor que tiene el promotor de polimerasa T7 y lleva preferiblemente una pequeña eliminación de por ejemplo, aproximadamente 10 a 25 nucleótidos con relación a la secuencia ARN objetivo nativa. La plantilla de ADN para la síntesis *in vitro* del competidor de ARN se sintetiza utilizando cebadores PCR que están entre 42 y 46 nucleótidos de longitud. En este ejemplo, los pares de cebador para la síntesis del ADN competidos CYTb 5 se muestran en la tabla 6.

TABLA 6

Cebadores Delantero e inverso utilizado para sintetizar la plantilla de ARN competidor para la medición de QC-RT-PCR de la expresión del gen CYTb5

cebador	cebador Competidor Delantero -	
Delantero	3740-179B	TAATACGACTCACTATAGGGAGGCACACCA CCCACGACGACTTGTG (SEQ. ID NO. 41)
cebador	cebador Competidor inverso -	
inverso	3740-179D	CTTCCGTGCTGAACGACTGCGAATCTTAGC GCCCTTCAAGTT (SEQ. ID NO. 42)

El cebador delantero se utiliza con el cebador inverso correspondiente para sintetizar la plantilla de ADN competidor. Los pares cebadores se combinan en una reacción PCR de polimerasa Taq Gold estándar de acuerdo con las condiciones recomendadas por el fabricante (Perkin-Elmer/Applied Biosystems, Foster City, CA). La mezcla de reacción contiene una concentración final de 250 nM de cada cebador y 10 ng de ADN cromosómico de *C. tropicalis* para la plantilla. La mezcla de reacción se coloca en un termociclizador durante 25 a 35 ciclos utilizando la mayor temperatura de hibridación posible durante las reacciones PCR para asegurar un producto PCR homogéneo (en este caso 62°C). Los productos PCR son purificados ya sea con gel o filtrados purificados para remover los nucleótidos no incorporados y cebadores. La plantilla de ADN competidor se cuantifica luego utilizando el método ($A_{260/280}$).

C. Síntesis del ARN competidor

El ADN de plantilla competidor se transcribe *in vitro* para hacer el ARN competidor utilizando el kit Megascript T7 De Ambion Biosciences (Ambion Inc., Austin, Texas). 250 nanogramos (ng) de plantilla de ADN competidor y los reactivos de transcripción *in vitro* se mezclan de acuerdo con las instrucciones proporcionadas por el fabricante. La mezcla de reacción se incuba durante 4 horas a 37°C. Las preparaciones de ARN resultantes se revisan luego mediante electroforesis de gel para las condiciones dado los mayores rendimientos y calidad del ARN competidor. Esto requiere frecuentemente optimización de acuerdo con las especificaciones del fabricante. La plantilla de ADN se remueve luego utilizando DNasa I como se describe en el kit Ambion. El competidor ARN se cuantifica luego mediante el método ($A_{260/280}$). Diluciones en serie del ARN (1 ng/ml a 1 femtograma (fg)/ml) se hacen para uso en las reacciones QC-RT-PCR y se almacenan la materia prima original a -70°C.

D. Reacciones QC-RT-PCR

Las reacciones QC-RT-PCR se desarrollan utilizando la polimerasa rTth de Perkin-Elmer(Perkin-Elmer/Applied Biosystems, Foster City, CA) de acuerdo con las condiciones recomendadas por el fabricante. La reacción de transcripción inversa se desarrolla en un volumen de 10 µl con una concentración final de 200 mM de cada uno de los dNTP, 1.25 unidades de Ruth de polimerasa, 1.0 mM de MnCl₂, 1X del amortiguador 10X suministrado con la enzima del fabricante, 100 ng de un ARN total aislado de un cultivo de crecimiento de fermentador de *C. tropicalis* y 1.25 mM del cebador inverso apropiado. Para cuantificar la expresión CYTb5 en *C. tropicalis* un cebador inverso apropiado es 5' CTTCCGTGCTGAACGACTGCG 3' (3740-179C). Varias mezclas de reacción se preparan para cada muestra de ARN caracterizada. Para cuantificar la expresión CYTb5 una serie de 8 a 12 de las mezclas de reacción QC-RT-PCR descritas anteriormente se dividen en alícuotas para diferentes tubos de reacción. Cada tubo de reacción de 1 ml de una dilución en serie contiene 100 pg a 100 fg del ARN competidor CYTb5 por ml que se agrega al llevar las mezclas de reacción final hasta el volumen final de 10 µl. Las mezclas de reacción QC-RT-PCR se mezclan y se incuban, a 70°C durante 15 min de acuerdo con los tiempos recomendados por el fabricante para que ocurra la transcripción inversa. En la terminación de la incubación de 15 minutos, la temperatura de la muestra se reduce a 4°C para detener la reacción y 40 µl de la mezcla de reacción PCR se agrega a la reacción para llevar el volumen total hasta 50 µl. La mezcla de reacción PCR consiste de una solución acuosa que contiene 0.3125 mM del cebador delantero 5' CACACCACCCAC GACGACTTGTG 3' (3740-179A), 3.125 mM MgCl₂ y 1X de amortiguador quelante suministrado con la enzima de Perkin-Elmer. Las mezclas de reacción se colocan en un termociclizador (Perkin-Elmer GeneAmp PCR System 2400, Perkin-Elmer/Applied Biosystems, Foster City, CA) y se desarrolla el siguiente ciclo PCR: 94°C durante 1 min. seguido por 94°C- durante 10 segundos seguido por 58°C durante 40 segundos durante 17 a 22 ciclos. La reacción PCR se completa con una incubación final a 58°C durante 2 min seguido por 4°C. En algunas reacciones en donde los productos PCR no se detectan se producen las muestras que se regresan al termociclizador para ciclos adicionales, este proceso se repite hasta que se produce suficientes productos PCR para cuantificar utilizando HPLC: El número de ciclos necesarios para producir suficiente producto PCR es una función de la cantidad del objetivo mRNA en los 100 ng de ARN celular total. En cultivos en donde el gen CYTb5 se expresa altamente hay suficiente mensaje CYTb5 mRNA presente y menos ciclos PCR (<17) que se requieren para producir cantidad cuantificable de producto PCR. Las concentraciones más bajas del mRNA objetivo presentan los mayores ciclos PCR que se requieren para producir una cantidad detectable de producto.

E. Cuantificación HPLC

Luego de la terminación de las reacciones QC-RT-PCR las muestras se analizan y cuantifican por HPLC. Cinco a quince microlitros de la mezcla de reacción QC-RT-PCR se inyectan en un HPLC 625 Bio-Compatible Waters con un detector ajustable 484 Waters. El detector se configure para medir una longitud de onda de 254 nm. El HPLC contiene una columna Sarasep marca DNasep™ (Sarasep, Inc., San Jose, CA) que se coloca dentro del horno y la temperatura se configura a 52°C. La columna se instala de acuerdo con la recomendación del fabricante de tener 30 cm de tubería PEEK calentada e instalada entre el inyector y la columna. El sistema se configure con una columna Sarasep marca Guard posicionada antes del inyector. Adicionalmente, hay un disco de filtro de 0.22 mm justo antes de la columna, dentro del horno. Se utilizan dos amortiguadores para crear un gradiente de elusión para redissolver y cuantificar los productos PCR a partir de las reacciones QC-RT-PCR. El amortiguador A consiste de 0.1M acetato de trietilamonio (TEAA) y 5% acetonitrilo (volumen a volumen). Amortiguador-B que consiste de 0.1 M TEAA y 25% acetonitrilo (volumen a volumen). Las muestras de QC-RT-PCR se inyectan en el HPLC y el gradiente lineal de 75% de amortiguador-A/25% amortiguador-B a 45% amortiguador-A/55% B que se corre durante 6 min a una tasa de flujo de 0.85 ml por minuto. El producto QC-RT-PCR del competidor de ARN que es más pequeño se eluye de la columna HPLC antes del producto QC-RT-PCR del CYTb5 mRNA (U). La cantidad de los productos QC-RT-PCR se grafican y se cuantifican con el módulo de datos 745 de Waters Corporation adjunto. Las proporciones log de la cantidad de producto CYTb5 mRNA QC-RT-PCR (U) a producto QC-RT-PCR competidor (C), según se mide por áreas pico, se grafican y la cantidad de ARN competidor requerido para igualar la cantidad de productos de CYTb5 mRNA determinado.

Ejemplo 12

25 Evaluación de nuevas Cepas en Frascos de Agitación

Las cepas CYTb5 y CPR amplificadas tales como las cepas HDC11-1 y HDC11-2 yH5343 se evalúan para la producción diácido en frascos de agitación. La producción diácido de las cepas CYTb5 y CPR también se compara con la producción diácido de varias cepas modificadas genéticamente de *C. tropicalis* 20336 [(HDC1-que contienen copias adicionales del gen CYP52A2A (Figuras 7A-7B); El HDC 10-2-contiene copias adicionales del gen CPRB gene (Figuras 4A-4B); HDC-15-contiene copias adicionales del gen CYP52A5A (Figuras 12A-12C); HDC16-2-contiene copias adicionales del gen CPRB y del gen CYP52A5A (Figuras 4A-4B y 12A-12C, respectivamente); y HDC23-3-contiene copias adicionales del gen CPRB y CYP52A2A (Figuras 4A-4B y 7A-7B, respectivamente) que se construyen de acuerdo con los procedimientos descritos en la patente U. S. No. 6,331,420 y WO 00/20566 A2.

35 A. Protocolo Estándar para Frasco de Agitación

Se inocula 100 ml de YEPD (Apéndice) con 1 ml de materia prima de glicerol preparada de *C. tropicalis* seguido por la adición de 1 gota de antiespumante SAG471 (dimetilpolisiloxano, disponible comercialmente de Osi Specialties, Sistersville, WV). Al cultivo se le permite crecer durante 20 hrs en un agitador a 30°C a 300 rpm. El cultivo de 100 ml se transfiere luego al medio DCA2 (Apéndice) mediante combinación;

100	ml	Cultivo-YEPD
100	ml	Glicerol [500g/L]
20	ml	YNB (Apéndice) [334 g/L]
780	ml	DCA2

La suspensión celular se divide en cinco frascos de agitación de 2000 ml seguido por la adición de una gota de SAG471. El cultivo se le permite crecer durante 30 hrs en un agitador a 30°C a 300 rpm. El medio se transfiere a botellas plásticas de autoclave y luego las células se centrifugan a 4068 g durante 5 min. El supernatante se descarta y las células se resuspenden en DCA3 (Apéndice):

100	ml	Glicerol	[500 g/L]
20	ml	YNB	[334 g/L]
880	ml	0.3 M	Amortiguador de potasio-fosfato pH=7.5

La suspensión celular se asigna a 500 ml en frascos de agitación en cantidad de 50 ml y 10% de w/v de octadecano (Sigma Chemical Company, St. Louis, MO) o se agrega 3% w/v ácido oleico (Fisher Scientific Co.). La bioconversión se permite tener lugar en una incubadora de agitación a 30°C y 300 μ m por 48 hrs. Posteriormente, los cultivos se

ES 2 298 276 T3

almacenan congelados a -20°C en el frasco de agitación. Antes del análisis los cultivos se descongelan y se calientan a 70°C , se mezclan y se transfieren a tubos Falcon de 50 ml. Las muestras se cosechan y las concentraciones de diácido se analizan mediante el procedimiento descrito adelante. LA Figura 18 describe el esquema utilizado para la extracción y el análisis de diácidos y inonoácidos de los caldos de cultivos de las fermentaciones.

B. Protocolo de la Concentración de Biomasa

1. Muestras de Ácidos Oléico

Las muestras se calientan a 70°C y se centrifugan para mezclar. Una alícuota de cada muestra (10 ml) se transfiere luego a un tubo Falcon de 50, seguido por la adición de 2.5 ml KOH (7%). Cada tubo Falcon se centrifuga durante 10 min a 3000 rpm y se descarta el supernadante. Luego se agrega agua (15 ml) a cada tubo Falcon que contiene células centrifugadas. Si el supernadante es claro luego de la adición del agua a las células centrifugadas, entonces las células se secan en un horno a 80°C hasta que dos pesos secos toman horas a parte que permanecen iguales. Si el supernadante es turbio luego de la adición del agua a las células centrifugadas, se agrega 2.5 ml KOH (7%) luego de la adición de KOH, las células centrifugadas de cada tubo se resuspenden al centrifugar, y recentrifugar durante 10 min a 3000 rpm seguido por remoción del supernadante. Luego las células se lavan con 15 ml de agua y se resuspenden mediante centrifugación. Las células se resuspenden en agua, se recentrifugan durante 10 min a 3000 rpm y el supernadante se descarta. Las células se resuspenden en un volumen pequeño de agua, se centrifugan y se vierten en botes de aluminio seco, repesados. Cada tubo Falcon se enjuaga luego 1-2 veces con agua en botes secos. Los botes secos se secan en un horno a 80°C hasta que dos pesos secos toman dos horas a parte que permanecen iguales.

2. Muestras de Octadecano

Las muestras se calientan a 70°C y se centrifugan para mezclar. Una alícuota A 10 ml se transfiere a un tubo Falcon de 50 ml, seguido por la adición de 2.5 ml KOH (7%). Luego se agrega hexano (5 ml) a cada tubo Falcon y se mezcla mediante centrifugación. Cada tubo Falcon que contiene muestra se centrifuga luego durante 10 min a 3000 rpm y el supernadante se descarta. Luego se agrega agua (15 ml) a las células centrifugadas que se resuspenden mediante centrifugación. Las células resuspendidas en cada tubo se centrifugan luego durante 10 min a 3000 rpm y luego se descarta el supernadante. Las células centrifugadas se lava con 15 ml de agua, se resuspenden en el agua a centrifugar y se centrifuga a 10 min a 3000 rpm. El supernadante se descarta y las células se resuspenden en un volumen pequeño de agua, se centrifugan y se vierten en botes de aluminio pre-pesados. Cada tubo Falcon se enjuaga 1-2 veces con agua en botes secos, y luego se secan en un horno a 80°C .

C. Protocolo de Extracción-Extracción de Caldo de Cultivo Utilizando Metil-t-butiléter

Las muestras se calientan a 70°C . Un frasco EPA con tapa se alquitrana sobre una balanza de dos puestos. Luego se pesa un caldo de cultivo de muestra en cada frasco y se registra el peso. La balanza se alquitrana y luego se pesa un estándar interno de un gramo (C15 Monoácido en KOH [10 g/kg]). El frasco se tapa y se agita para mezclar. Las etapas mencionadas anteriormente se repiten para todas las muestras, usualmente se corren en grupos de 6 a la vez. Posteriormente se le remueve la tapa a los frascos. Se pipetea HCl (6N, 6.8 ml) en cada frasco que se tapa y se agita para mezclar. Luego se agrega Metil-t-butiléter (20 ml) a cada frasco, que se tapa y se centrifuga durante 1 min. Luego se pesa Sulfato de magnesio Anhidro (MgSO_4 , 5 g) en un plato para peso. Cada frasco se agita luego y se agrega inmediatamente MgSO_4 a cada frasco de agitación. Cada frasco se tapa herméticamente y se le da una agitación rápida. Posteriormente, la tapa en cada frasco se afloja cuidadosamente para aliviar la presión. La tapa en cada frasco se apreta y el frasco se centrifuga durante 1 min. Las etapas de la adición de MgSO_4 para centrifugar el frasco tapado se repite para cada frasco. Los frascos se enfrían luego durante aproximadamente 5 min y se centrifugan durante 1 min. Los contenidos de los frascos se les permite reposa. Un frasco de 4 ml se etiqueta para cada muestra y 1 ml de solución de cada frasco EPA se pipetea en cada frasco de 4 ml etipetado.

D. Metilación de Material Extractado

Trifluoruro de Boro (1 ml) en reactivo de methanol ($\text{BF}_3\text{-MeOH}$) se agrega a cada frasco de 4 ml. Los frascos se tapan luego y se agitan para mezcla. Posteriormente, los frascos se colocan en un bloque seco de 50°C y se dejan en el bloque seco durante un total de 20 min. Los frascos luego se remueven del bloque y se enfrían durante 5 min, seguido por la adición de 0.5 ml de hexano a cada frasco. Luego de la adición de una solución de NaCl saturado (5 ml) a cada frasco (71.5 g NaCl en 200 ml de agua destilada), los frascos se tapan luego y se agitan durante 1 min. Cada frasco se le permite reposar hasta que se separan las capas. Posteriormente, se etiqueta un frasco para cromatografía de gas (GC) para cada muestra. La mayor parte de la capa superior del frasco de 4 ml se transfiere luego al frasco GC1. La capa inferior del frasco de 4 ml no se transfiere al frasco GC. Cada frasco GC se tapa luego cuidadosamente y se coloca en la bandeja de automuestreado GC. Una vez se cargan todas las muestras del día, se inicia el GC. Durante una serie, usualmente dos blancos de methanol se emplean utilizando el método almacenado como TEMP.MET en el integrador al inicio del día para preparar el GC. Un blanco de methanol se corre después de 10 muestras y como la última muestra mantiene la columna GC limpia. Las condiciones GC son como siguen:

- HP-Innowax 30 m por 32 mm ID columna con 0.5 m de espesor de película
- Proporción de división 1:100

ES 2 298 276 T3

- Presión de cabeza de columna 13.5 psig
- Temperatura de inyector 240°C
- Temperatura de detector FID 250°C
- Programa de temperatura de horno:
 - 90°C rampa a 190°C a 7°C/min luego rampa 250°C a 12°C/min y se mantiene durante 30 min.

Esto se guarda en integradores como DCAME2.MET.

Los resultados de los frascos se muestran adelante en las 7-10.

TABLA 7

Bioconversión de 3% w/v de ácido oléico mediante diferentes cepas recombinantes de Candida tropicalis

<u>Experimento 1</u>					
	Genes amplificados	Promedio C18:1 Diácido (g/kg)	No. de Experimentos	Desviación estándar (%)	% de mejora
H5343	---	17.5	3	7.4	0.0
HDC1	CYP52A2A	16.5	3	17.5	-5.4
HDC10-2	CPR B	17.5	3	17.2	0.3
HDC15	CYP52A5A	21.5	3	16	23.3
HDC23-3	CPR B + CYP52A2A	23.3	3	8.7	33.5

TABLA 8

Bioconversión de 3% p/v de ácido oléico mediante cepas recombinantes de Candida tropicalis

<u>Experimento 2</u>					
Cepa	Genes amplificados	Promedio C18:1 Diácido (g/kg)	No. de Experimentos	Desviación estándar (%)	% de mejora
H5343	--	13.7	2	7.7	0.0
HDC11-1	CPR B + Cyt BS	17.3	2	5.7	25.7
HDC11-2	CPR B + Cyt B5	21.0	2	2.4	53.2
HDC16-2	CPR B + CYP52A5A	16.0	2	8.4	16.3
HDC23-3	CPR B + CYP52A2A	23.3	2	3.9	70.1

TABLA 9

Bioconversión de octadecano mediante cepas recombinantes de Candida tropicalis

Experimento 1 y 2					
Cepa	Genes amplificados	Producto Total Mono y diácidos (g/kg)	% producto total mejorado	Producto Total (Normalizado) % de mejoramiento	% total del producto. Mejoramiento
H5343	--	6.6/5.2	0.0/0.0	0/0	0
HDC11-1	CPR B + CYTb5	15.2/10.1	130.6/93.5	74.3/39.1	56.7±17.6
BDC11-2	CPR B + CYTb5	18.2/17	175.8/225.7	100/94.4	97.2±2.8
HDC16-2	CPR B + CYP52A5A	17.1/15.2	160.1/191.2	91.1/79.9	85.5±5.6
HDC23-3	CPR B + CYP52A2A	8.5/17.7	29.8/239.1	16.9/100	58.5±41.5

TABLA 10

Bioconversión de octadecano mediante diferentes cepas recombinantes de Candida tropicalis

Cepa	Genes amplificados	Producto Total	Producto total
		Mono y diácidos (g/kg)	% mejoramiento
H5343	--	11.2	0.0
HDC10-2	CPR B	22.1	98.0
HDC11-2	CPRB + CYTb5	16.1	44.3
HDC15	CYP52A5A	13.0	16.5
HDC16-2	CPRB + CYP52A5A	19.4	73.8

Estos datos demuestran que la producción diácido se incrementa en cepas *C. tropicalis* modificadas genéticamente HDC11- y HDC11-2 que contiene un número de copia incrementada de genes CY7b5/CPRB comparado con la cepa tipo silvestre H5343. La producción de diácido también se incrementa en las cepas *C. tropicalis* modificadas genéticamente HDC10-2, HDC15, HDC16-2, y HDC23-3.

Ejemplo 13

*Clonación y caracterización de genes CPR y b5 citocromo C. tropicalis 20336 (CYTb5)*A. Clonación del gen CYTb5 de *C. tropicalis* 20336

El gen CYTb5 se aísla de la tercera librería genómica utilizando un fragmento PCR como una sonda. La sonda de fragmento PCR para CYTb5 se genera después de amplificación PCR de ADN genómico de *Saccharomyces cerevisiae* con cebadores de oligonucleótido que se designan para amplificar una región que utiliza el gen CY7b5 disponible de *S. cerevisiae* del Centro Nacional para Información de Biotecnología. Un cebador delantero 3698-66A.

El 5' ATAAGAATCCGGCCGCTGAACGAGCCAGAACCCACATCCACATCCAGGAG 3' y a cebador inverso 3698-66B 5' CCTTAATrAAGGATAACCACATCcATACGTCGC 3' se hacen basado en la secuencia de *S. cerevisiae* GYTb5. Estos cebadores se utilizan en combinaciones en forma de pares en una reacción de PCR con ADN polimerasa Taq (Perkin-Elmer Cetus, Foster City, CA) de acuerdo con el procedimiento recomendado por el fabricante. Un producto PCR de aproximadamente 1036 bp se obtiene. Este producto se purifica de gel azarosa utilizando Qiaquick (Qiagen, Chatsworth, CA) y se liga al vector pCR2.1™ (LA Figura 19. Invitrogen LaJolla, CA) de acuerdo con las recomendaciones del fabricante. Este fragmento PCR se utiliza como una sonda en el aislamiento del homólogo.

go *C. tropicalis* 20336 CYTb5. La tercer librería genómica se selecciona utilizando esta sonda CYTb5 y un clon que contiene el gen CYTb5 se obtiene de longitud completa. El clon contiene un gen que tiene regiones de codificación de proteína y regulación (LAS Figuras 1A-1 C). Un cuadro de lectura abierta de 387 nucleótidos codifican una proteína CYTb5 de 129 aminoácidos (Figuras 1A-1C).

B. Confirmación de Función del gen CYTb5 clonado

El CYP51, lanosterol 14 α -desmetilasa, es el sitio de acción de la mayoría de categorías de agentes antifúngicos denominados DMI para inhibidores de desmetilasa. Agentes terapéuticos tales como ketoconazole (Kc), un agente antifúngico imidazol N-sustituido proporciona un ensayo conveniente para la actividad CYP51 basado en las concentraciones inhibitorias mínimas (MIC) como se describe en Vanden Bossche *et al.*, Antimicrobial Agents Chemother. 17:922-928, 1980. Las cepas *S. cerevisiae* cpr deficientes en reductasa (cpr) P450 citocromo han mostrado ser hipersensibles al Kc, debido al la reducción en el flujo de electrones para CYP5T, el objetivo de Kc como se describe en Sutter y Loper, *supra*. La recuperación de esta hipersensibilidad se logra al clonar genes basados en la transformación de una librería genómica tipo silvestre en un vector multicopia dentro de una cepa cpr1 como se describe en Truan *et al.*, *supra*. Los genes responsables para la recuperación de esta sensibilidad al Kc se muestran por ser aquellos que codifican el CYTb5 citocromo, que muestran que en este sistema, CYTb5 es capaz de imitar la funcionalidad CPR cuando está presente en altos números de copia. La inhibición del aldosterol desmetilasa mediante una cierta concentración de Kc resulta en la radiación de crecimiento de *S. cerevisiae*. En la ausencia de CPR, la inhibición de crecimiento es más pronunciada y ocurre en una concentración mas baja de Kc. En la ausencia de gen CPR, el gen CYTb5 puede suprimir el fenotipo a server como el donante de electrones para CYP51, a pesar de la ineficiencia, y alivia parcialmente el efecto supresor de crecimiento de Kc. Por lo tanto, el *S. cerevisiae* recombinante que contiene un *C. tropicalis* CYTb5 será capaz de crecer en un medio que contiene niveles intermedios de Kc. Este ensayo se utiliza para confirmar que los genes *C. tropicalis* clonados codifican para CYTb5.

El gen CYTb5 se clona bajo el control de un promotor GAL presente en el vector de expresión de *S. cerevisiae* que contiene CpYE5 que se introduce en el mutante *S. cerevisiae* DC10 que contiene un gen CPR interrumpido y la cepa recombinante se coloca en placas en un medio mínimo que contiene diferentes concentraciones de Kc. La cepa DC10 que contiene el gen CPR del *S. cerevisiae* tipo silvestre (en el plásmido pTS20) y el gen CY7b5 de *C. tropicalis* (en el plásmido pYES2b5) es capaz de crecer en medio mínimo que contiene 0.078 mg/ml de Kc. De otra parte, el DC10 que lleva el pYES2.0 sin ningún inserto es capaz de crecer en concentraciones de 0.039 mg/ml de Kc. Las cepas recombinantes se prueban en cultivos líquidos para confirmar el resultado de las placas. Mientras que el DC10 contiene pYES2.0 sin ningún inserto es capaz de crecer en 0.0024 mg/ml de Kc, el DC10 que expresa b5 del promotor GAL es resistente a 0.0098 mg/ml de Kc. Esto confirma que el gen clonado codifica para la CYTb5.

C. Clonación y caracterización de los genes CPR de *C. tropicalis* 20336

Para clonar el CPR se utiliza una sonda heteróloga basada en la secuencia de ADN para la secuencia del gen CPR de *C. tropicalis* 750 para aislar el gen CPR *C. tropicalis* 20336

1) Clonación del alelo CPRA

Aproximadamente 25,000 partículas fago de la primera librería genómica de *C. tropicalis* 20336 se seleccionan con un fragmento de 1.9 kb Bam HI - Nde I del plásmido pCU3RED (Ver Picattagio *et al.*, Bio/Technology 10:894-898 (1992), que contiene la mayoría del gen CPR *C. tropicalis* 750. Cinco clones que hibridan la sonda se aíslan y el plásmido de ADN de estos clones lambda se rescata y caracterizan por análisis de enzima de restricción. El análisis de enzima de restricción sugiere que todos los cinco clones son idénticos pero no es claro que está presente un gen CPR completo.

El análisis PCR se utilice para determinar si un gen GPR completo está presente en cualquiera de los cinco clones. Los cebadores degenerados se preparan para regiones altamente conservadas de genes CPR conocidos (Ver Sutter *et al*, J. Biol. Chem. 265:16428-1636,1990) (LA Figura 21). Se sintetizan dos cebadores de grasa en la región de unión FMN (FMN1 y FMN2). Un cebador se sintetiza para la región de unión PAD (FAD), y un cebador para la región de unión NADPH (NADPH) (Table 4). Esos cuatro cebadores se utilizan en experimentos de amplificación PCR utilizando como un aislado de ADN de plásmido plantilla cuatro de los cinco clones descritos anteriormente. Los cebadores FMN y FAD sirven como cebadores delanteros y el cebador NADPH como el cebador inverso de las reacciones PCR. Cuando se utilizan diferentes combinaciones de cebadores inversos y delantero, no se obtienen productos PCR de cualquiera de los plásmidos. Sin embargo, todos los productos de tamaño esperado amplificado de combinaciones de cebador con un plásmido que contiene el gen CPR *C. tropicalis* 750 (control positivo). Probablemente la mayor razón para la falla de los pares de cebadores para amplificar un producto, es que todos los cuatro clones contienen un gen CPR truncado. Uno de los cuatro clones (pHKM1) se secuencia utilizando los cebadores Triplex 5' y el Triplex 3' (Table 4) cuyo flanco en el inserto y en el sitio de clonación múltiple en el vector de clonación y el cebador degenerado basado en el sitio de unión NADPH descrito anteriormente. El cebador NADPH falla en producir cualquier dato de secuencia y esto es consistente con el análisis PCR. Las secuencias obtenidas con los cebadores Triplex se comparan con la secuencia CPR de *C. tropicalis* 750 utilizando el programa Mac vectorTM (Oxford Molecular Group, Campbell, CA). La secuencia obtenida con el cebador Triplex 3' mostró similitud con una secuencia interna del gen CPR *C. tropicalis* 750 que confirma que el pHKM1 contiene una versión truncada de un gen CPR 20336, el pHKM1 tiene un inserto de 3.8 kb el cual incluye una región de codificación 12 kb del gen CPR acompañado por 2.5 kb de ADN

ascendente (Figura 22). Aproximadamente 0.85 kb del gen CPR 20336 CPR que codifica la porción C-terminal de la proteína CPR se pierde de este clon.

En razón de que la primera librería Clontech produce un gen CPR truncado, la segunda librería prepara por Clontech se selecciona para aislar un gen CPR de longitud completa. Se obtiene tres clones putativos CPR. Los tres clones, que tienen insertos en el rango de 5-7 kb, se designan pHKM2, pHKM3 y pHKM4. Todos los tres se caracterizan por PCR utilizando los cebadores degenerados descritos anteriormente. Los productos pHKM2 y pHKM4 dan productos PCR con dos conjuntos de cebadores internos. El pHKM3 da un producto PCR con los cebadores FAD y NADPH lo que sugiere que este clon probablemente contiene un gen CPR truncado. Los tres plásmidos se secuencian parcialmente utilizando los dos cebadores Triples y un tercer cebador cuya secuencia se selecciona de la secuencia de ADN cercada al extremo truncado del gen pHKM1. Este análisis confirma que el pHKM2 y 4 tienen secuencias que transapan el pHKM1 y que ambas contienen la región 3' de la región del gen CPR que desaparece de pHKM1. Las porciones de los insertos pHKM1 y pHKM4 se secuencian y el gen CPR de longitud completa se identifica. Con base en la secuencia de ADN y análisis PCR se concluye que el pHKM1 contiene el promotor putativo de la región y 1.2 kb de la secuencia que codifica una porción (5' end) de un gen CPR. El pHKM4 tiene 1.1 kb de AND que translapa el pBKM1 y contiene el restante (3' and) de un gen CPR junto con la región no traducida descendente (Figura 23). Ambos de estos dos plásmidos contienen un gen CPRA completo con una región promotora en la dirección 5'. El CPRA es de 4206 nucleótidos de longitud e incluye una región reguladora y una región de codificación de proteína (definida por los nucleótidos 1006-3042) que es de 2037 bases pares de longitud y codifican para una proteína putativa de 679 aminoácidos (Figuras 2A-2B y 3A-3C, respectivamente). La proteína CPRA, cuando se analiza mediante el programa de alineación de proteína del paquete software GeneWorks™ (Oxford Molecular Group, Campbell, CA), mostró homología extensiva a proteínas CPR a *C. tropicalis* 750 y *C. maltosa*.

2) Clonación del alelo CPRB

Para clonar el Segundo alelo CPR, la tercera librería genómica, preparada por Henkel, se selecciona utilizando fragmentos de ADN de pHKM1 y pHKM4 utilizados como sondas. Se obtienen cinco clones y estos se secuencian con los tres cebadores internos utilizados para secuencias CPRA. Estos cebadores se designan PRK1.F3, PRK1.F5 y PRK4.R20 (Table 4) y dos cebadores externos (M13-20 y T3 [Stratagene]) para la región de poliligador presente en el vector de clonación pBK-CMV. Los análisis de secuencias sugieren que cuatro de estos clones, designados pHKM5 to 8, contienen insertos que son idénticos al alelo CPRA aislado anterior. Todos los cuatro parecen contener un gen CPR de longitud completa. El quinto clon es muy similar al alelo CPRA, especialmente en la región de cuadro de lectura abierta en donde la identidad es muy alta. Sin embargo, hubo diferencias muy significativas en las regiones no traducidas 5' y 3'. Esto sugiere que el quinto clon es el alelo para el CPRA. El plásmido se designa pHKM9 (Figura 24) y una región de 4.14 kb de este plásmido se secuencia y el análisis de esta secuencia confirma la presencia del alelo CPRB, que incluye una región reguladora y una región de codificación de proteína (definida por nucleótidos 1033-3069) como se muestra en las Figuras 4A-4B. La secuencia de aminoácido de la proteína CPRB se muestra en las Figuras 5A-5C.

Ejemplo 14

Identificación del Gen CYTb5 Inducido por Ácidos Grasos Seleccionados y Substratos Alcano

Los genes cuya transcripción se enciende por la presencia de ácidos grasos seleccionados o sustratos alcano se han identificado usando el ensayo QC-RT-PCR. Este ensayo se utiliza para medir la expresión del gen CYTb5 en cultivos de crecimiento en fermentador de *C. tropicalis* ATCC 20962. Este método involucra el aislamiento de ARN celular total de cultivos de *C. tropicalis* y la cuantificación de un mRNA muy específico dentro de aquella muestra a través del diseño y uso de secuencia específica, cebadores QC-RT-PCR y un competidor ARN. La cuantificación se logra a través del uso de concentraciones conocidas de competidor ARN homólogos en las reacciones QC-RT-PCR. El cADN amplificado por QC-RT-PCR resultante se separa y se cuantifica a través del uso de fase inversa de HPLC de fase inversa de pares de ión. Este ensayo se utiliza para caracterizar la expresión del gen CYTb5 de *C. tropicalis* ATCC 20962 en respuesta al ácido graso y sustrato de alcano. Los genes que se incluyen se identifican mediante el cálculo de su concentración de mRNA en varios momentos antes y después de inducción. La Figura 25 suministra un ejemplo de cómo la concentración de mRNA para CYTb5 se puede calcular utilizando el ensayo QC-RT-PCR. La proporción log desconocido (U) a producto competidor (C) se grafica versus la concentración de ARN del competidor presente en las reacciones QC-RT-PCR. La concentración del competidor que resulta en una proporción log de U/C de cero, representa el punto en donde la concentración de ARN mensajero desconocida es igual a la concentración del competidor. La Figura 25 permite el cálculo de la cantidad de mensajes CYTb5 presente en 100 ng de ARN total aislado de muestras celulares tomadas en 0, 0.5, 1, 2, y 3 horas después de la adición de Emersol® en una serie de fermentador. A partir de este análisis, es posible determinar la concentración de CYTb5 mRNA presente en 100 mg de ARN celular total. En la gráfica contenida en la LA Figura 25 esta toma 113.75 pg de competidor para igualar el número de mARNs de CYTb5 en 100 ng de ARN aislado de células justo antes (tiempo 0) de la adición del sustrato, Emersol®. En muestras celulares tomadas en 0.5, 1, 2, y 3 horas después de la adición de Emersol® esta toma 293.69, 356.01, 372.24, y 264.61 pg de competidor ARN, respectivamente. Los resultados obtenidos durante esta serie de fermentación según se mide por QC-RT-PCR también se expresa como inducción relativa del gen CYTb5 como se muestra en la tabla 11 adelante.

Emersol[®]

Estos datos demuestran que el CYTb5 se induce más de 3.0 veces dentro de una hora después de la adición de Emersol[®]. Este análisis demuestra claramente que la expresión CYTb5 en *C. tropicalis* 20962 se puede inducir por la adición de Emersol[®] al medio de crecimiento.

Ejemplo 15

10 Integración de los genes CYTb5 y CPRB seleccionados dentro del Genoma de *Candida tropicalis*

Con el fin de integrar los genes seleccionados dentro del cromosoma de *C. tropicalis* 20336 o sus descendientes, hay una secuencia de ADN objetivo, que puede o no ser un gen intacto, dentro del cual se pueden insertar los genes. También debe haber un método para seleccionar el evento de integración. En algunos casos la secuencia de ADN objetivo y el marcador seleccionable son los mismos y, si es así, entonces debe haber también un método para obtener uno del gen objetivo como un marcador seleccionable siguiendo el evento de integración. En *C. tropicalis* y sus descendientes, un gen que se ajusta con estos criterios es el URA3A, que codifica orotidina-5'-fosfato descarboxilasa. Utilizando este como un objetivo para integración, las variantes URA⁻ de *C. tropicalis* se pueden transformar en tal una forma para regenerar un genotipo URA⁺ por vía de recombinación homóloga (Figura 26). Dependiendo del diseño del vector de integración, se pueden integrar uno o más genes dentro del genoma al mismo tiempo. Utilizando un gen URA3A de división orientado como se muestra en la Figura 27, la integración homóloga produciría por lo menos una copia del gen de interés que se inserta entre las porciones divididas del gen URA3A. Más aún, debido a la alta similitud de secuencia entre los genes URA3A y URA3B, la integración del constructor puede ocurrir en ambos lugares URA3A y URA3B. Posteriormente, un oligonucleótido diseñado con una eliminación en una porción del gen URA basado en la secuencia idéntica a través de los genes URA3A y URA3B, se puede utilizar para producir los transformantes *C. tropicalis* que son de nuevo una vez URA⁻ pero que todavía llevan uno o más genes de elección integrados nuevamente (Figura 26). Las variantes URA⁻ de *C. tropicalis* también se pueden aislar por vía de otros métodos tal como mutagénesis clásica o mutación espontánea. Utilizando bien los protocolos establecidos, la selección de las cepas URA⁻ se puede facilitar mediante el uso de ácido 5-fluoroorotico (5-FOA) como se describe, por ejemplo, en Boeke *et al.*, Mol. Gen. Genet. 197:345-346 (1984). La utilidad de este enfoque para la manipulación de este punto *C. tropicalis* se ha documentado bien como se describe, por ejemplo, en Picataggio *et al.*, Mol. and Cell. Biol. 11:4333-4339 (1991); Rohrer *et al.*, Appl. Microbiol. Biotechnol. 36:650-654 (1992); Picataggio *et al.*, Biol/Technology 10:894-898 (1992); patente U.S. No. 5,648,247; patente U.S. No. 5,620,878; patente U.S. No. 5,204,252; patente U.S. No. 5,254,466.

35 A. Construcción, de un vector de integración URA, pURAIin.

Se diseñan cebadores y se sintetiza con base en la secuencia 1712 bp del gen URA3A de *C. tropicalis* 20336 (ver Figura 28), cebador URA3A Set #1a y #1b (Tabla 4) en PCR con cADN genómico de *C. tropicalis* 20336 para amplificar las secuencias URA3A entre nucleótidos 733 y 1688 como se muestra en la Figura 28. Los cebadores se diseñan para introducir los sitios de restricción 5' Asc I y 3' Pac I dentro del fragmento URA3A amplificado resultante. Los sitios Asc I y Pac I se escogen debido a que estos sitios no están presentes dentro del gen CYTb5 o CPRB. Se utiliza el cebador URA3A Set #2 en PCR de ADN genómico de *C. tropicalis* 20336 como una plantilla, para amplificar las secuencias URA3A entre el nucleótido 9 y 758 como se muestra en la Figura 27. Se diseña el cebador URA3A set #2a y #2b (Tabla 4) para introducir los sitios de restricción 5' Pac I y 3' Pme I únicos en el fragmento URA3A resultante. El sitio Pme I tampoco está dentro de los genes CYTb5 y CPRB. Los fragmentos PCR del gen URA3A se purifican, se restringen con enzimas de restricción Asc I, Pac I y Pme I y se ligan a un gel purificado, se digieren con Asc I - Pme I limpiado con QuiaexII de plásmido pNEB 193 (Figura 29) comprando de New England Biolabs (Beverly, MA). La ligación se desarrolla con un número equimolar de Terminal ADN en 16°C durante 16 hr, usando ligasa de ADN T4 (New England Biolabs). Las ligaciones se transforman en células *E. coli* XL1-Blue (Stratagene, LaJolla, CA) de acuerdo con las recomendaciones del fabricante. Se aíslan colonias blancas, se hacen crecer, se aísla el ADN de plásmido y se digiere con Asc I - Pme I para confirmar la inserción del URA3A modificado en el pNEB193. El vector de integración base resultante se denomina PURAIin (Figura 30).

55 B. Amplificación de CYTb5 y CPRB de ADN genómico de *C. tropicalis* 20336

El gen que codifica el CYTb5 se amplifica por vía de PCR usando cebadores (cebadores CYT b5#1, GGGTTAAT-TAACATACTTCAAGCAGTTTGG; y el cebador CYT b5 # 2, CCCTTAATTAAGGGGGGATGGAAGTGGCCG) para introducir los sitios de clonación Pac I. Ver tabla 4. Estos cebadores se diseñan con base en la secuencia de ADN determinada para CYTb5. El kit UEL Itma (Perkin Elmer Cetus, Foster City, CA) se utilice de acuerdo con las especificaciones de los fabricantes. El producto de amplificación CYTb5 PCR es de 1998 pares bases de longitud, produciendo 1088 bp de ADN en la dirección 5' del codón de inicio CYTb5 y 486 bp en la dirección 3' del codón de detención para el CYTb5 ORF. Con el fin de generar un producto PCR libre de error para clonación, un fragmento Bgl/II interno de un PCR se desplaza con un fragmento Bgl/II de un PCR diferente que no contiene mutaciones en la región. Un producto CYT b5 libre de error final se genera que contiene sitios de clonación sensibles Pac I.

El gen que codifica el CPRB de *C. tropicalis* 20336 se amplifica del ADN genómico por vía de PCR utilizando los cebadores (CPR B#1 y CPR B#2) con base en la secuencia de ADN determinado para CPRB (Figuras 4A-4B). Estos cebadores se diseñan para introducir sitios de clonación únicos de Pac I. El kit Expand Hi-Fi Taq PCR kit

(Boehringer Mannheim, Indianapolis, IN) se utiliza de acuerdo con las especificaciones de los fabricantes: El producto PCR CPRB es 3266 bp en longitud, que produce 747 bp pf de ADN en la dirección 5' del codón de inicio CPRB y 493 bp en la dirección 3' del codón de parada para el CPRB ORF. Los productos PCR resultantes se aíslan por vía de electroforesis de gel de azarosa, se purifican utilizando QiaexII y se digieren con Pac I. Los fragmentos de PCR se purifican, desalinizan y se concentran utilizando un Microcon 100 (Amicon, Beverly, MA).

Los procedimientos de amplificación anteriormente descritos también se aplican al gen CPRA (Figuras 2A-2B) utilizando los cebadores indicados respectivamente (Tabla 4).

C. Clonación del gen CPRB en pURAIin

La siguiente etapa es clonar el gen CPRB en el vector de integración pURAIin. En un aspecto preferido de la presente invención, ningún ADN externo diferente de aquel proporcionado específicamente mediante las secuencias de sitio de restricción sintética se incorporan en el ADN que se clona en el genoma *C. tropicalis*, es decir, con la excepción del sitio de restricción de ADN solo las secuencias de ADN *C. tropicalis* nativas se incorporan en el genoma. El pURAIin se digiere con Pac I, Qiaex II limpio, y se desfosforila con Fosfatasa alcalina de Camarón (SAP) (United States Biochemical, Cleveland, OH) de acuerdo a las recomendaciones del fabricante. Aproximadamente 500 ng de Pac I linearizado pURAIin fue desfosforilado durante 1 hr a 37°C utilizando SAP, a una concentración de 0.2 Unidades de enzima por 1 pmol de ADN Terminal. La reacción se detiene mediante inactivación por 65°C durante 20 min.

Antes de su uso, el fragmento CPRB Pac I se deriva utilizando los cebadores mostrados en la tabla 4 que se secuencia y se compara con CPRB para confirmar que el PCR no introduce cambios de pares base de ADN que pudieran resultar en cambios de aminoácidos. Luego de la confirmación, el CPRB se liga al plásmido pURAIin que también se ha digerido con Pac I. El Pac I digerido con pURAIin es desfosforilado, y se liga con el producto expand CPR Hi-Fi PCR como se describió previamente. La mezcla de ligación se transforma en *E. coli* XL1 Blue MRF' (Stratagene) y se seleccionan varias colonias resistentes y se seleccionan para construcciones correctas que pudieran contener secuencia de vector, el gen URA3A invertido, y el gen CPRB amplificado (Figuras 4A-4B) de 20336. La digestión de Asc I- Pme I le confirma una construcción exitosa, pURAREDBin.

En una forma similar, a la anterior, el gen CPRA descrito aquí se clona en el pURAIin. El fragmento Pac I del gen CPRA, cuya secuencia de nucleótido se da en las Figuras 2A-2B, es derivable por métodos conocidos por aquellos expertos en la técnica.

1) Construcción de Vectores Utilizado para Generar HDC11

El vector de integración previamente construido que contiene CPRB, pURAREDBin, se selecciona como el vector de inicio. Este vector se digiere parcialmente con Pac I y el fragmento linearizado se aísla con gel. El Pac I activo se destruye por tratamiento con polimerasa AND T4 y el vector se religa. El posterior aislamiento y digestión completa de este nuevo plásmido produce un vector que contiene ahora un solo sitio Pac I activo. Este fragmento es aislado con gel, desfosforilado y ligado con el fragmento Pac I CYTb5. Los vectores que contienen los genes CYTb5 y CPRB orientados en las direcciones opuestas (conectados en los extremos 5'), pURAIin CPR b5 O, se generan.

D. Confirmación de Integración CYTb5 en el Genoma de *C. tropicalis*

Basado en la construcción del pURAIin CPR b5 O, utilizado para transformar el H5343 ura-, un esquema para detectar la integración se idea. El ADN genómico de transformantes se digiere con Dra III que es una enzima que corta dentro de los genes URA3A, URA3B, CPRB y CYTb5. la digestión del ADN genómico en donde ha ocurrido una integración en el lugar URA3A se esperaría que resulte en la elucidación de un fragmento de 3.6 Kb luego de hibridación Southern con una sonda CPRB. Las hibridaciones Southern de esta digestión con fragmentos del CPRB se utilizan para selección de estos eventos de integración. La intensidad de esta señal de banda del software se utiliza como una medición del número de eventos de integración (es decir la mayoría de copias del gen CYTb5 que están presentes, la más fuerte es la señal de hibridación).

Los prototrofos URA transformados por *C. tropicalis* H5343 se hacen crecer a 30°C, 170 rpm, en 10 ml de medio SC-uracilo para preparación de ADN genómico. El ADN genómico se aísla mediante los métodos descritos previamente. El ADN genómico se digiere con Dra III. Un gel de 0.95% de azarosa se utiliza para preparar una transferencia de hibridación Southern. El ADN del gel se transfiere a una membrana de filtro de nylon MagnaCharge (MSI Technologies, Westboro, MA) de acuerdo con el método de transferencia de Sambrook *et al.*, *supra*. Para la hibridación Southern se utilice un fragmento de AND CPRB de 3.3 Kb como una sonda de hibridación. 300 ng de AND CPRB se etiquetan utilizando un sistema de detección y etiquetado usando ECL Direct (Amersham) y el Southern se procesa de acuerdo con las especificaciones del kit ECL. La transferencia se procesa en un volumen de 30 ml de fluido de hibridación que corresponde a 0.125 ml/cm². Luego de una prehibridación a 42°C durante 1 hr, 300 ng de sonda CPRB se agrega y la hibridación continua durante 16 hr a 42°C. Luego de la hibridación, las transferencias se lavan dos veces durante 20 min cada uno a 42°C en urea de lavado continuo. Dos lavados secundarios de 5 min se conducen seguido por detección de acuerdo con instrucciones. Las transferencias se exponen durante 16 horas (hr) como se recomienda.

ES 2 298 276 T3

La integración se confirma por la detección de un fragmento Dra III 3.6 Kb del ADN genómico de los transformantes pero no con el control *C. tropicalis* 20336. Las cepas CYTb5 y CPRB integradas resultantes se denominan HDC11. Las variantes de esta cepa se designan arbitrariamente 11-1 a 11-4 basado en el número relativo de eventos de integración que ocurren.

Cuando se comparan con el H5343 pariente, HDC 11-1, HDC11-3 y HDC11-4 cada uno contiene una copia adicional única del gen CYTb5 y HDC 11-2 que contiene dos copias adicionales del gen CYTb5.

Apéndice

Composición Media	Agua Destilada	1,000 ml
Caldo de cultivo LB	agarosa superior NZCYM	
Triptón Bacto	10 g	Extracto de levadura Bacto 10 g
Extracto de levadura Bacto	5 g	Ácidos casamino Bacto 1 g
Cloruro de sodio	10 g	Extracto de levadura Bacto 5 g
Agua destilada	1,000 ml	Cloruro de Sodio 5 g
		Sulfato de magnesio 0.98 g
Agar LB r	(anhidro)	
Triptón Bacto	10 g	Agarosa 7 g
Extracto de levadura Bacto	5 g	Agua destilada 1,000 ml
Cloruro de sodio	10 g	
Agar	15 g	Caldo YEPD
Agua destilada	1,000 ml	Extracto de levadura Bacto 10 g
		Bacto Peptona 20 g
Agarosa superior LB		Glucosa 20 g
Triptón Bacto	10 g	Agua destilada 1,000 ml
Extracto de levadura Bacto	5 g	
Cloruro de sodio	10 g	Agar YEPD r*
Agar	7 g	Extracto de levadura Bacto 10 g
Agua destilada	1,000 ml	Bacto Peptona 20 g
		Glucosa 20 g
Caldo NZCYM		Agar 20 g
Digestión de caseína Bacto	10 g	Agua destilada 1,000 ml
Ácidos casamino Bacto	1 g	
Extracto de levadura Bacto	5 g	YNB
Cloruro de Sodio	5 g	Extracto de levadura 3 g/L
Sulfato de magnesio	0.98 g	Maltosa 3 g/L
(hidratado)		Pepton 5 g/L

ES 2 298 276 T3

	Agua destilada	1,000ml	Dextrose	10 g/L
	Agar NZCYM			
5	Digestión de caseína Bacto	10g		
	Ácidos Casamino Bacto	1 g		
	Extracto de levadura	5 g		
10	Cloruro de sodio	5 g		
	Sulfato de Magnesio (hidratado)	0.98 g		
15	Agar	15 g		
	Medio DCA2	g/l	SC - uracil*	
		3.0	Base de levadura nitrógeno sin aminoácidos	6.7g
20	Peptona			
	Extracto de levadura	6.0	Glucosa	20g
	Acetato de sodio	3.0	Bacto-agar	20g
25	Base de levadura nitrógeno (Difco)	6.7	Mezcla en gotas	2g
	Glucosa (anhidro)	50.0	Agua destilada	1,000ml
	Fosfato potasio (dibásico, trihidrato)	7.2		
30	Fosfato potasio (monobásico, anhidro)	9.3		
	Medio DCA3	g/l		
35	0.3 M Amortiguador de fosfato que contiene pH	7.5		
	Glicerol	50		
40	Base de Nitrógeno levadura Difco)	6.7		
	Mezcla en gotas			
45	Adenina	0.5g	Alanina	2g
	Arginina	2g	Asparagina	2g
	Ácido aspártico	2g	Cisteína	2g
50	Glutamina	2g	ácido glutámico	2g
	Glicina	2g	Histidina	2g
	Inositol	2g	Isoleucina	2g
55	Leucina	10g	Lisina	2g
	Metionina	2g	para-ácido Aminobenzoico	0.2g
	Fenilalanina	2g	Prolina	2g
60	Serina	2g	Treonina	2g
	Triptófano	2g	Tirosina	2g
	Valina	2g		
65	*ver Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press, USA (1994), incorporada aquí como referencia.			

REIVINDICACIONES

1. Una célula anfitriona transformada con un ácido nucleico que codifica la proteína CYTb5 (citocromo b5) que tiene la secuencia de aminoácido establecida en la SEQ. ID. NO. 2 en donde todos los anfitriones contienen una interrupción de la ruta β -oxidación

2. Una célula anfitriona de acuerdo con la reivindicación 1 en donde la célula anfitriona es una célula de levadura.

3. Una célula anfitriona de acuerdo con la reivindicación 2 en donde la célula de levadura se selecciona del grupo que consiste de *Yarrowia*, *Candida*, *Bebaromices*, *Saccharomyces*, *Schizosaccharomyces*, y *Pichia*.

4. Una célula anfitriona de acuerdo con la reivindicación 3 en donde la célula anfitriona cándida se selecciona del grupo *C. tropicalis*, *C. maltosa*, *C. apicola*, *C. paratropicalis*, *C. albicans*, *C. cloacae*, *C. guilliermondii*, *C. intermedia*, *C. lipolytica*, *C. parapsilosis*, y *C. zeylenoides*.

5. Una célula anfitriona de acuerdo con la reivindicación 1 en donde el ácido nucleico está contenido en un vector de expresión.

6. Una célula anfitriona de acuerdo con la reivindicación 5 en donde el vector de expresión se selecciona del grupo que consiste de plásmido, fagémido, fago, cósmido, cromosoma artificial de levadura, vector de ADN lineal y vector de integración.

7. Una célula anfitriona de acuerdo con la reivindicación 6 en donde el plásmido se selecciona del grupo que consiste plásmido episómico de levadura y plásmido replicante de levadura.

8. Una célula anfitriona de acuerdo con la reivindicación 1 en donde el ácido nucleico que codifica la proteína CYTb5 comprende la SEQ. ID. NO. 1.

9. Una célula anfitriona de acuerdo con la reivindicación 1 en donde el ácido nucleico que codifica la proteína CYTb5 comprende los residuos de ácido nucleico numerados 1109 a 1495 de la SEQ. ID. NO. 1.

10. Una célula anfitriona de acuerdo con la reivindicación 1 en donde la célula anfitriona contiene una o más copias de un gen seleccionado del grupo que consiste del gen CPR (P450 RED; citocromo P450 reductasa), gen CYP (P450 ALK; citocromo P450 monooxigenasa) y combinaciones de estos.

11. Un método para incrementar la producción de ácido dicarboxílico que comprende:

Proporcionar una célula anfitriona que tiene ninguno, uno o más genes CYTb5;

Incrementar, en la célula anfitriona, el número de genes CYTb5 que codifican una proteína CYTb5 que tienen la secuencia de aminoácidos establecida en la SEQ. ID. NO. 2 en donde todos los anfitriones contienen una interrupción de la ruta de β -oxidación; y

Cultivar la célula anfitriona en un medio que contiene un sustrato orgánico que induce la expresión del gen CYTb5 para efectuar la producción incrementada de ácido dicarboxílico.

12. Un método para incrementar la producción de un ácido dicarboxílico de acuerdo con reivindicación 11 en donde la célula anfitriona es una célula de levadura.

ES 2 298 276 T3

```

      10      20      30      40      50      60
ACATACTTCA AGCAGTTTGG CGACATAGTG AACCTCAAGT TATCACGGAA CAAGACGACG

      70      80      90     100     110     120
GGCAAGAGCA AGCACTACGG GTTTATAGAG TTCACGTCGC CTGAAGTTGC CCAGATCGCG

      130     140     150     160     170     180
GCGGAGACGA TGAACAATA CTGTGTGTTT GGACACTTGA TCAAATGTGA GGTGTGCAGC

      190     200     210     220     230     240
GAGCCGTTCA AGGACTTGTT CAAGGACTCG AAGAGGAAGT TCAAGGTGAT TCCCTGGAAG

      250     260     270     280     290     300
AAGATCGCGA AGGATAAGCA CGATAAGCCA AAGTCCGCGA AGGAGTGGGC GAAGTTGGTG

      310     320     330     340     350     360
GAGAAGTTCC AAGAGTCCAA GAAGAAGAAG CAGGAGGAGT TGAAGAGTAA AGGTATTGAT

      370     380     390     400     410     420
TTTGATTTGG CTGCTATATA AAGGAGATAA GAGAGGAGGA TGACAAGCCG AAACGAGCAT

      430     440     450     460     470     480
TCTGTTGATG TGTAAAGCAG GTATAGATAA TAGCGGATAA CGTAAATAA GAGATCTCCA

      490     500     510     520     530     540
ACTTCCAAC TCCAACCTCC GACCCCTATC TTTTGGGCGA GAGGGATTGG TATGTAGTGG

      550     560     570     580     590     600
TGAGGGAGAG GAGGATATTT TGTTTTGGCT AATTGGGATA AATTATCCCA GTCAGTTGAA

      610     620     630     640     650     660
AGAGCGAGGC GTAAGCCATT TCTTTTCTTA ACTGCAAATA GCATACAGAT GCGATAGTTA

      670     680     690     700     710     720
ACGAGAGAG AAATCAAGAG CAGGTGACTA CATACATAGA TAGTGACATT ATAATAACAT

      730     740     750     760     770     780
GGCGCATCAT TGGTCTATG TAGCTGGCAG GGTATTATC AACCTGAAT AGTTTAATAA

      790     800     810     820     830     840
AAATCGTACC ATGAATGTAT GCATAGAAGC AATAAGGAAG CCTGTGCCTG TGAGTAGTAG

      850     860     870     880     890     900
CAGTAGCGGG GGGAGACGCT AGTTTAGGGG TAAATGTCA GCACATGAAC ACCAGTTGAA

      910     920     930     940     950     960
GTGGGTGCCA ATCAAGTAAG AACATCTTGT GAAAAATCAA AAGCAATGGT ATATGTGTTC

      970     980     990    1000    1010    1020
CTGCATACAG TGCTGGAGTC AACGAGCCAA AAAAAAAAAA GAAAGAAAGA GAGAAAAACT

      1030    1040    1050    1060    1070    1080
TATCGTATAA AAACACACA AAAATTTCCC AATCCCAATT CCTTCATTCT TCTTCTTTTA

      1090    1100    1110    1120    1130
CTGATTTAAC CCACAGATAC ATACAATT ATG ACC GAC ACA GAC ACC ACG ACC ACC
              M   T   D   T   D   T   T   T   T>

      1140    1150    1160    1170    1180
ATC TAC ACC CAC GAA GAG GTT GCC CAG CAC ACC ACC CAC GAC GAC TTG
I   Y   T   H   E   E   V   A   Q   H   T   T   H   D   D   L>

```

FIG. 1A

ES 2 298 276 T3

```

1190      1200      1210      1220      1230
TGG GTT ATT CTC AAT GGT AAG GTC TAC AAC ATC TCC AAC TAT ATA GAC
W V I L N G K V Y N I S N Y I D>

      1240      1250      1260      1270
GAG CAC CCA GGT GGT GAA GAA GTC ATT CTT GAT TGC GCC GGC ACA GAC
E H P G G E E V I L D C A G T D>

1280      1290      1300      1310      1320
GCC ACT GAA GCC TTT GAC GAC ATT GCC CAC TCC GAC GAG GCC CAC GAG
A T E A F D D I G H S D E A H E>

1330      1340      1350      1360      1370
ATC TTG GAA AAG TTG TAC ATT GGT AAC TTG AAG GGC GCT AAG ATT GTT
I L E K L Y I G N L K G A K I V>

      1380      1390      1400      1410      1420
GAG GCC AAG CAC GCG GAG TCG TTC AGC ACG GAA GAA GAC TCG GGT ATC
E A K H A Q S F S T E E D S G I>

      1430      1440      1450      1460      1470
AAC TTC CCA TTG ATT GCT GTT GGT GTG TTT TTG GCT GCT TTC GGT GTC
N F P L I A V G V F L A A F G V>

      1480      1490      1500      1510      1520
TAC TAC TAC AAG ACC AAC TTT GCC TAAGG ATAACAGCA GTACAGTTGA
Y Y Y K T N F A>

      1530      1540      1550      1560      1570      1580
AGGACAGGGT AGAGGAGATG AGAAAAAAGC GGAACCCAC AAAGATTATT TTCACACATC

      1590      1600      1610      1620      1630      1640
ACATGGAGGG GCTGATCCCA CTTTTTGACG TCAATATCCA CAGCAGGAAG AAAGAAAGAA

      1650      1660      1670      1680      1690      1700
AGAAAGAAAG TCTATGGAAG AGGAAATGGA TCACATTAGA GCTTTTCTTT ATGTAACATA

      1710      1720      1730      1740      1750      1760
TATATATATA TAACTAATA CAGATTACA GATACACGAC ATGACCGCAG GGCTTATCAT

      1770      1780      1790      1800      1810      1820
CTGATGGTGC CCAAAAAAAA AAATCCACTG TGGATGAGCC TAGTTAGGAG ATATCGGAGT

      1830      1840      1850      1860      1870      1880
AGCTCATTCT TTGATATCT AGGTCTTCCT CTCTGGATT CTACGTTGGT ACTTGGTGCT

      1890      1900      1910      1920      1930      1940
ACACGATGAG ATCACCAGGT GTCAATCTGG AGTTTGGTGG AAAGTGIGTT GATTTTTTTA

      1950      1960      1970      1980      1990      2000
GTAAGCAAGA ATTTGTTGAG TTCTATTGGA TGTCTCTGGT CGGCCACTTC CATCCCCCA

      2010      2020      2030      2040      2050      2060
CCCCCTGTCT TGTCTGTCT TGTCTTATT TTTTGGGTCT GTTGGCGGAA GTAAGACGCA

      2070      2080      2090      2100      2110      2120
CGCACAGGAG GAGCACGACG GATAAATATC CACTTTTTTC ACACGCGTCT ATTGACGGCT

      2130      2140      2150      2160      2170      2180
TGTGTGAATT GTGGGAATA CGGATAAGGG GGTATACCAC ACACACACAT ATCTAACATA

      2190      2200      2210      2220      2230      2240
TCAGACCACT TTCTATAAGA GATCTCATGA TCCCCTTGAG AGTTGATGCA AGTCTATGCT

```

FIG. 1B

ES 2 298 276 T3

```

2250      2260      2270      2280      2290      2300
CCTGTGATAT TGCCCCCCCC CCCCCAAGGA AGGGCGGGGC ATGTTATCAG GGACCTGGAT

2310      2320      2330      2340      2350      2360
GAACCCCTGA TGGCGGTGTG AGTAGATGCA AGAGAGGTTG TGCTTTGGAA GTAGCTGAAG

2370      2380      2390      2400      2410      2420
GTGTAGGGAC ATCCGGTAGT ATAGTTCTCT TGAAGGATCA TGCCAGCTCC CTTTCTGTGG

2430      2440      2450      2460      2470      2480
CTCTCTGGAA GCTCTGCATC TTCTCTTCTG TGAACAGCG TGGAGTTACG AAAGGTACCC

2490      2500      2510      2520      2530      2540
TGTGGTGAGT TCAACAAGA CATGGCTCTA CAAGCTGTCG AGGATAAAG TAATTAAACA

2550      2560      2570      2580      2590      2600
ACATGTATAT ATATTAAATA ACGGATCCGT GGTGCTAGAT TGTGGTAGAT GTTTAGTATC

2610      2620      2630      2640      2650      2660
GTTTATCACC TCTAGTGAAA ACTAGCATTI GATTCCATTA GTCATCAGTA GTTGATGTTA

2670      2680      2690      2700      2710
CATTCAACCA AATGAAGGTC GGTCCAAGAT CCAAGAATT CAAAAGCTT

```

FIG. 1C

```

catcaagatc atctatgggg ataattacga cagcaacatt gcagaaagag cgttgggtcac 60
aatcgaaaga gcctatggcg ttgccgtcgt tgaggcaaat gacagcacca acaataacga 120
tggtcccagt gaagagcctt cagaacagtc cattgttgac gcttaaggca cggataatta 180
cgtggggcaa aggaacgcgg aattagttat ggggggatca aaagcggaaag atttgtgttg 240
cttgtgggtt ttttccttta tttttcatat gatttctttg cgcaagtaac atgtgccaat 300
ttagtttgtg attagcgtgc cccacaattg gcacgtgga cgggcgtgtt ttgtcatacc 360
ccaagtotta actagctcca cagtctcgac ggtgtctcga cgatgtcttb ttccaccctt 420
cccatgaatc attcaaagt gttgggggat ctccaccaag ggcaccggag ttaatgctta 480
tgtttctctc actttgggtg tgattggggt agtctagtga gttggagatt ttcttttttt 540
cgcaggtgtc tccgatatcg aaatttgatg aatatagaga gaagccagat cagcacagta 600
gattgcottt gtagttagag atgttgaaca gcaactagtt gaattacacg ccaccacttg 660
acagcaagtg cagtgaagtc taaacgatgc agccagagtg tcaccacca aagtagcaca 720
gtggagttgt tgttgttgtt gttggcaggg ccatattgct aaacgaagac ctgacgttg 780
aaaccaagc ttaagaacaa aaataaaaaa aattcatacg acaattccaa agccattgat 840
ttacataatc aacagtaaga cagaaaaaac tttcaacatt tcaaagttcc ctttttctta 900
ttacttcttt tttttcttct tctctcttct ccttctgttt ttcttacttt atcagtcttt 960
taottgtttt tgcaattcct catcctctct ctactctctc tcaccatggc tttagacaag 1020
ttagatttgt atgtcatcat aacattgggt gtgcgtgtag ccgcctattt tgctaagaac 1080
cagttccttg atcagcccca ggacaccggg ttctcaaca cggacagcgg aagcaactcc 1140
agagacgtct tgctgacatt gaagaagaat aataaaaaca cgttgttgtt gtttgggtcc 1200
cagacgggtg cggcagaaga ttacgccaac aaattgtcca gagaattgda ctccagattt 1260
ggcttgaaaa cgatggttgc agatttcgct gattacgatt gggataactt cggagatatt 1320
accgaagaca tcttgggtgt ttcatttgtt gccacctatg gtgagggtga acctaccgat 1380
aatgccgacg agttccacac ctggttgact gaagaagctg acactttgag taccttgaaa 1440
tacaccgtgt tcgggttggg taactccaag tacgagttct tcaatgccat tggtagaaag 1500
tttgacagat tgttgagcga gaaaggtggg gacaggtttg ctgaatacgc tgaagggtgat 1560
gacggtactg gcaccttga cgaagatttc atggcctgga aggacaatgt ctttgaogcc 1620
ttgaagaatg atttgaactt tgaagaaaag gaattgaagt acgaaccaaa cgtgaaattg 1680
actgagagag acgacttgtc tgetgtgtac tcccaagttt ccttgggtga cccaacaag 1740
aagtacatca actccgaggg catcgacttg accaagggtc cattcgacca ccccaccca 1800
tacttggcca gaatcaccca gacgagagag ttgttcagct ccaaggacag acactgtatc 1860
cacgttgaat ttgacatttc tgaatcgaaac ttgaaatata ccaccggtga ccatctagct 1920
atctggccat ccaactcoga cgaaaacatt aagcaatttg ccaagtgtt cggattggaa 1980
gataaactcg aactgttat tgaattgaag gcgttggact ccacttacac catcccattc 2040
ccaaccccaa ttacctacgg tgctgtcatt agacaccatt tagaaatctc cgttccagtc 2100
tcgagacaat tctttttgtc aattgotggg ttgtctctg atgaagaaac aaagaaggct 2160
tttaccagac ttggtgggtga caagcaagaa ttccgcgcca aggtcaccgg cagaaagttc 2220
aacattgcg atgccttgtt atattcctcc aacaacgctc catggtccga tgttcttttt 2280
gaattcctta ttgaaaacgt tccacacttg actccagtt actactccat ttogtctctg 2340
tcattgagtg aaaagcaact catcaacgtt actgcagttg ttgaagccga agaagaagct 2400
gatggcagac cagtcactgg tgttgtcacc aacttgttga agaacgttga aattgtgcaa 2460
aacaagactg gcgaaaagcc acttgtccac tacgatttga gcggcccaag aggcaagttc 2520
aacaagttca agttgccagt gcatgtgaga agatccaact ttaagttgcc aaagaactcc 2580
accacccag ttatcttgat tggtcaggt actggtgttg cccatttgag aggttttgtc 2640
agagaaagag ttcaacaagt caagaatggg gtcaatgttg gcaagacttt gttgttttat 2700
ggttgacaga actccaacga ggactttttg tacaagcaag aatgggcccga gtacgttctt 2760
gttttgggtg aaaactttga gatgttcaat gccttctcca gacaagaccc atccaagaag 2820
gtttacgtcc aggataagat tttagaaaac agccaacttg tgcaogagtt gttgactgaa 2880
ggtgccatta tctacgtctg tggatgtgcc agtagaatgg ctagagacgt gcagaccaca 2940
atttccaaga ttgttgctaa aagcagagaa attagtgaag acaaggctgc tgaattggtc 3000
aagtcttggg aggtccaaaa tagataccaa gaagatgttt ggtagactca aacgaatctc 3060

```

FIG. 2A

ES 2 298 276 T3

```

tctttctccc aacgcattta tgaatcttta ttctcattga agctttacat atgtttctaca 3120
ctttattttt tttttttttt ttattattat attacgaaac ataggtcaac tatatatact 3180
tgattaaatg ttatagaaac aataactatt atctaactgt ctacttcttt ggcatcgaca 3240
tcaacattac cgttcccatt accgttgcgg ttggcaatgc cgggatattt agtacagtat 3300
ctccaatdcg gatttgagct attgtagatc agctgcaagt cattctccac cttcaaccag 3360
tactttatact tcatctttga cttcaagtcc aagtcataaa tattacaagt tagcaagaac 3420
ttctggccat ccacgatata gacgttattc acgttattat gcgacgtatg gatgtggtta 3480
tccptattga acttctcaaa cttcaaaaac aaccccacgt ccgcacaagt cattatcaac 3540
gacaagttct ggctcacgtc gtccggagctc gtcaagttct caattagatc gttcttggtta 3600
ttgatcttct ggtactttct caattgctgg aacacattgt cctcgttggt caaatagatc 3660
ttgaacaact ttttcaacgg gatcaacttc tcaatctggg ccaagatctc cgcggggatc 3720
ttcagaaaca agtcctgcaa cccctggctg atgggtctccg ggtacaacaa gtccaagggg 3780
cagaagtgtc taggcacgtg tttcaactgg ttcaacgaac atgttcgaca gtagttcgag 3840
ttatagttat cgtacaacca ttttggtttg atttcgaaaa tgacggagct gatgccatca 3900
ttctcctggg tctcttcata gtacaactgg cacttcttcg agaggctcaa ttctcctgtag 3960
ttcccggtcca agatattcgg caacaagagc ccgtacogct cacggagcat caagtctgtg 4020
ccctggttgt tcaacttggt gatgaagtcc gaggtcaaga caatcaactg gatgtcgatg 4080
atctgggtgog ggaacaagtt ctgcatttt agctcgatga agtcgtacaa ctcacacgtc 4140
gagatatact cctgttcctc cttcaagagc cggatccgca agagcttggt cttcaagtag 4200
tcgttg
4206

```

FIG. 2B

ES 2 298 276 T3

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

FIG. 3A

ES 2 298 276 T3

Lys Gly Pro Phe Asp His Thr His Pro Tyr Leu Ala Arg Ile Thr Glu
 260 265 270
 Thr Arg Glu Leu Phe Ser Ser Lys Asp Arg His Cys Ile His Val Glu
 275 280 285
 Phe Asp Ile Ser Glu Ser Asn Leu Lys Tyr Thr Thr Gly Asp His Leu
 290 295 300
 Ala Ile Trp Pro Ser Asn Ser Asp Glu Asn Ile Lys Gln Phe Ala Lys
 305 310 315 320
 Cys Phe Gly Leu Glu Asp Lys Leu Asp Thr Val Ile Glu Leu Lys Ala
 325 330 335
 Leu Asp Ser Thr Tyr Thr Ile Pro Phe Pro Thr Pro Ile Thr Tyr Gly
 340 345 350
 Ala Val Ile Arg His His Leu Glu Ile Ser Gly Pro Val Ser Arg Gln
 355 360 365
 Phe Phe Leu Ser Ile Ala Gly Phe Ala Pro Asp Glu Glu Thr Lys Lys
 370 375 380
 Ala Phe Thr Arg Leu Gly Gly Asp Lys Gln Glu Phe Ala Ala Lys Val
 385 390 395 400
 Thr Arg Arg Lys Phe Asn Ile Ala Asp Ala Leu Leu Tyr Ser Ser Asn
 405 410 415
 Asn Ala Pro Trp Ser Asp Val Pro Phe Glu Phe Leu Ile Glu Asn Val
 420 425 430
 Pro His Leu Thr Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Ser Leu Ser
 435 440 445
 Glu Lys Gln Leu Ile Asn Val Thr Ala Val Val Glu Ala Glu Glu Glu
 450 455 460
 Ala Asp Gly Arg Pro Val Thr Gly Val Val Thr Asn Leu Leu Lys Asn
 465 470 475 480
 Val Glu Ile Val Gln Asn Lys Thr Gly Glu Lys Pro Leu Val His Tyr
 485 490 495
 Asp Leu Ser Gly Pro Arg Gly Lys Phe Asn Lys Phe Lys Leu Pro Val
 500 505 510
 His Val Arg Arg Ser Asn Phe Lys Leu Pro Lys Asn Ser Thr Thr Pro
 515 520 525

FIG. 3B

ES 2 298 276 T3

Val	Ile	Leu	Ile	Gly	Pro	Gly	Thr	Gly	Val	Ala	Pro	Leu	Arg	Gly	Phe
530						535					540				
Val	Arg	Glu	Arg	Val	Gln	Gln	Val	Lys	Asn	Gly	Val	Asn	Val	Gly	Lys
545					550					555					560
Thr	Leu	Leu	Phe	Tyr	Gly	Cys	Arg	Asn	Ser	Asn	Glu	Asp	Phe	Leu	Tyr
			565						570					575	
Lys	Gln	Glu	Trp	Ala	Glu	Tyr	Ala	Ser	Val	Leu	Gly	Glu	Asn	Phe	Glu
			580					585					590		
Met	Phe	Asn	Ala	Phe	Ser	Arg	Gln	Asp	Pro	Ser	Lys	Lys	Val	Tyr	Val
		595					600					605			
Gln	Asp	Lys	Ile	Leu	Glu	Asn	Ser	Gln	Leu	Val	His	Glu	Leu	Leu	Thr
610						615					620				
Glu	Gly	Ala	Ile	Ile	Tyr	Val	Cys	Gly	Asp	Ala	Ser	Arg	Met	Ala	Arg
625					630					635					640
Asp	Val	Gln	Thr	Thr	Ile	Ser	Lys	Ile	Val	Ala	Lys	Ser	Arg	Glu	Ile
			645						650					655	
Ser	Glu	Asp	Lys	Ala	Ala	Glu	Leu	Val	Lys	Ser	Trp	Lys	Val	Gln	Asn
		660						665					670		
Arg	Tyr	Gln	Glu	Asp	Val	Trp									
675															

FIG. 3C

tatatgatat	atgatatatc	ttcctgtgta	attattattc	gtattcgtta	atacttacta	60
catttttttt	tctttattta	tgaagaaaag	gagagtccgt	aagttgagtt	gagtagaata	120
ggctgttgtg	catacgggga	gcagaggaga	gtatccgacg	aggaggaaact	gggtgaaatt	180
tcatctatgc	tgttgcgtec	tgtactgtac	tgtaaatcct	agatttccta	gagggtgttc	240
tagcaataaa	agtgtttcaa	gatacaattt	tacaggcaag	ggtaaaggat	caactgatta	300
gcggaagatt	ggtgttgcc	gtggggttct	tttatttttc	atatgatttc	tttgccgag	360
taacatgtgc	caatctagtt	tatgattagc	gtacctccac	aattggcctc	ttggacgggc	420
gtgttttgtc	ttaccccaag	ccttatttag	ttccacagtc	tcgacgggtg	ctgcgcgag	480
tcttctccca	ccctcgcag	gaatcattcg	aagttgttgg	gggatctcct	ccgcagttta	540
tgttcatgtc	tttcccactt	tgggtgtgat	tggggtagcg	tagtgagttg	gtgattttct	600
tttttcgcag	gtgtctccga	tatcgaagtt	tgatgaatat	aggagccaga	tcagcatggt	660
atattgcctt	tgtagataga	gatgttgaa	aacaactagc	tgaattacac	accacgccta	720
aacgatgccc	acaggggtgc	accgccaact	gcagttgggt	ggagtgtgtg	ttggcagggc	780
catattgcta	aacgaagaga	agtagcaca	aaccacaagg	taagaacaat	taaaaaaatt	840
catacgacaa	ttctacagcc	atttacataa	tcaacagcga	caaatggagac	agaaaaaact	900
ttcaacattt	caaagttccc	tttttcctat	tacttctttt	tttcttctct	tcctttcatt	960
tcctttcctt	ctgcttttat	tactttacca	gtcttttgct	tgtttttgca	attcctcctc	1020
ctcctcctca	ccatggcttt	agacaagtta	gatttgtatg	tcatacatac	attgggtggc	1080
gctgtggccg	cctatttttg	taagaaccag	ttccttgatc	agccccagga	caccgggttc	1140
ctcaacacgg	acagcggaag	caactccaga	gaogtcttgc	tgacattgaa	gaagaataat	1200
aaaaacacgt	tgttgttgtt	tgggtcccag	accggtacgg	cagaagatta	cgccaacaaa	1260
ttgtcaagag	aattgcactc	cagatttggc	ttgaaaacca	tggttgcaga	tttcgctgat	1320
tacgattggg	ataacttcgg	agatatcact	gaagatatct	tgggtgtttt	catcggtgcc	1380
acctacgggtg	agggtgaacc	taccgacaat	gccgacgagt	tcacacacctg	gttgactgaa	1440
gtagctgaca	ctttgagtag	tttgagatat	accgtgttog	ggttgggtaa	ctccacctac	1500
gagttcttca	atgctattgg	tagaaagttt	gacagattgt	tgagtgagaa	agggtggtgac	1560
agatttgctg	aatatgctga	agggtgacgac	ggcactggca	ccttggaagg	agatttcatg	1620
gcctggaagg	ataatgtcct	tgacgccttg	aagaatgact	tgaactttga	agaaaaggaa	1680
ttgaagtacg	aaccaaacgt	gaaattgact	gagagagatg	acttgtctgc	tgccgactcc	1740
caagtttctt	tgggtgagcc	aaacaagaag	tacatcaact	ccgagggcat	cgacttgacc	1800
aagggtccat	togaaccacac	ccaccatac	ttggccagga	tcaccgagac	cagagagttg	1860
ttcagctcca	aggaaagaca	ctgtattcac	gttgaatttg	acatttctga	atcgaaactg	1920
aaatacacca	ccggtgacca	tctagccatc	tggccatcca	actccgacga	aaacatcaag	1980
caatttgcca	agtgtttcgg	attggaagat	aaactcgaca	ctgttattga	attgaaggca	2040
ttggaactca	cttacacccat	tccattccca	actccaatta	cttacgggtgc	tgtcattaga	2100
caccatttag	aaatctccgg	tccagtctcg	agacaattct	ttttgtcgat	tgtgtgggtt	2160
gtcctctgatg	aagaaacaaa	gaagactttc	accagacttg	gtggtgacaa	acaagaattc	2220
gccaccaagg	ttacccgcag	aaagttcaac	attgcccgatg	ccttggtata	ttcctccaac	2280
aacactccat	ggtccgatgt	tctttttgag	ttccttattg	aaaacatcca	acacttgact	2340
ccacgttact	actccatttc	ttcttcgtcg	ttgagtgaag	aacaactcat	caatgttact	2400
gcagtcgttg	aggccgaaga	agaagccgat	ggcagaccag	tcactgggtg	tgttaccac	2460
ttgttgaaag	acattgaaat	tgcgcaaaac	aagactggcg	aaaagccact	tgttccactac	2520
gatttgagcg	gcccgaagag	caagttcaac	aagttcaagt	tgccagtgc	cgtgagaaga	2580
tccaacttta	agttgccaaa	gaactccacc	accccagtta	tcttgattgg	tccaggtact	2640
ggtgttgccc	cattgagagg	tttcgttaga	gaaagagttc	aacaagtcaa	gaatgggtgc	2700
aatgtttggc	agactttggt	gttttatggg	tgcagaaact	ccaaagagga	ctttttgtac	2760
aagcaagaat	gggcccagta	cgcttctgtt	ttgggtgaaa	actttgagat	gttcaatgcc	2820
ttctctagac	aagaccatc	caagaagggt	tacgtccagg	ataagatttt	agaaaacagc	2880
caacttggtg	acgaattggt	gaccgaagg	gccattatct	acgtctgtgg	tgacgccagt	2940

FIG. 4A

```

agaatggcca gagacgtcca gaccacgato tccaagattg ttgccaaaag cagagaaatc 3000
agtgaagaca aggcgcgtga attgggtcaag tccctggaaag tccaaaatag ataccaagaa 3060
gatgtttggt agactcaaac gaatctctct ttctcccaac gcatttatga atattctcat 3120
tgaagtttta catatgttct atatttcatt ttttttttat tatattacga aacataggtc 3180
aactatatat acttgattaa atgttataga aacaataatt attatctact cgtctacttc 3240
tttggcattg gcattggcat tggcattggc attgcogttg cggttggtaa tgcggggata 3300
tttagtacag tatctccaat ccggatttga gctattgtaa atcagctgca agtcattctc 3360
caccttcaac cagtacttat acttcatctt tgacttcaag tccaagtcac aaatattaca 3420
agttagcaag aacttctggc catccacaat atagacgtta ttcacgttat tatgcgacgt 3480
atggatatgg ttatccttat tgaacttctc aaacttcaaa aacaacccca cgtcccgcaa 3540
cgtcattatc aacgacaagt tctgactcac gtcgtcggag ctcgtaagt tctcaattag 3600
atcgttcttg ttattgatct tctggtaact tctcaactgc tggaaacacat tgcctctggt 3660
gttcaaatag atcttgaaca acttcttcaa gggaaatcaac ttttcgatct gggccaagat 3720
ttccgccggg atcttcagaa acaagtcctg caaccctgg tcgatgggtct cgggggtacaa 3780
caagtctaag gggcagaagt gtctaggcac gtgtttcaac tggttcaagg aacatgttcg 3840
acagtagtgc gagttatagt tatcgtaaaa ccactttggc ttgatttcga aaatgacgga 3900
gctgatccca tcattctcct ggttcctttc atagtacaac tggcatttct tcgagagact 3960
caactcctcg tagttcccg ccaagatatt cggcaacaag agcccgtagc gctcacggag 4020
catcaagtcg tggccctggt tgttcaactt gttgatgaag tccgatgtca agacaatcaa 4080
ctggatgtcg atgatctggt gcggaacaaa gttcttgcac tttagctcga tgaagtcgta 4140
caact 4145

```

FIG. 4B

ES 2 298 276 T3

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

FIG. 5A

ES 2 298 276 T3

Lys Gly Pro Phe Asp His Thr His Pro Tyr Leu Ala Arg Ile Thr Glu	260	265	270
Thr Arg Glu Leu Phe Ser Ser Lys Glu Arg His Cys Ile His Val Glu	275	280	285
Phe Asp Ile Ser Glu Ser Asn Leu Lys Tyr Thr Thr Gly Asp His Leu	290	295	300
Ala Ile Trp Pro Ser Asn Ser Asp Glu Asn Ile Lys Gln Phe Ala Lys	305	310	315
Cys Phe Gly Leu Glu Asp Lys Leu Asp Thr Val Ile Glu Leu Lys Ala	325	330	335
Leu Asp Ser Thr Tyr Thr Ile Pro Phe Pro Thr Pro Ile Thr Tyr Gly	340	345	350
Ala Val Ile Arg His His Leu Glu Ile Ser Gly Pro Val Ser Arg Gln	355	360	365
Phe Phe Leu Ser Ile Ala Gly Phe Ala Pro Asp Glu Glu Thr Lys Lys	370	375	380
Thr Phe Thr Arg Leu Gly Gly Asp Lys Gln Glu Phe Ala Thr Lys Val	385	390	395
Thr Arg Arg Lys Phe Asn Ile Ala Asp Ala Leu Leu Tyr Ser Ser Asn	405	410	415
Asn Thr Pro Trp Ser Asp Val Pro Phe Glu Phe Leu Ile Glu Asn Ile	420	425	430
Gln His Leu Thr Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Ser Leu Ser	435	440	445
Glu Lys Gln Leu Ile Asn Val Thr Ala Val Val Glu Ala Glu Glu Glu	450	455	460
Ala Asp Gly Arg Pro Val Thr Gly Val Val Thr Asn Leu Leu Lys Asn	465	470	475
Ile Glu Ile Ala Gln Asn Lys Thr Gly Glu Lys Pro Leu Val His Tyr	485	490	495
Asp Leu Ser Gly Pro Arg Gly Lys Phe Asn Lys Phe Lys Leu Pro Val	500	505	510
His Val Arg Arg Ser Asn Phe Lys Leu Pro Lys Asn Ser Thr Thr Pro	515	520	525

FIG. 5B

ES 2 298 276 T3

Val	Ile	Leu	Ile	Gly	Pro	Gly	Thr	Gly	Val	Ala	Pro	Leu	Arg	Gly	Phe	
530						535					540					
Val	Arg	Glu	Arg	Val	Gln	Gln	Val	Lys	Asn	Gly	Val	Asn	Val	Gly	Lys	
545					550					555					560	
Thr	Leu	Leu	Phe	Tyr	Gly	Cys	Arg	Asn	Ser	Asn	Glu	Asp	Phe	Leu	Tyr	
			565						570					575		
Lys	Gln	Glu	Trp	Ala	Glu	Tyr	Ala	Ser	Val	Leu	Gly	Glu	Asn	Phe	Glu	
			580					585					590			
Met	Phe	Asn	Ala	Phe	Ser	Arg	Gln	Asp	Pro	Ser	Lys	Lys	Val	Tyr	Val	
		595					600					605				
Gln	Asp	Lys	Ile	Leu	Glu	Asn	Ser	Gln	Leu	Val	His	Glu	Leu	Leu	Thr	
610						615					620					
Glu	Gly	Ala	Ile	Ile	Tyr	Val	Cys	Gly	Asp	Ala	Ser	Arg	Met	Ala	Arg	
625					630					635					640	
Asp	Val	Gln	Thr	Thr	Ile	Ser	Lys	Ile	Val	Ala	Lys	Ser	Arg	Glu	Ile	
			645						650					655		
Ser	Glu	Asp	Lys	Ala	Ala	Glu	Leu	Val	Lys	Ser	Trp	Lys	Val	Gln	Asn	
			660					665					670			
Arg	Tyr	Gln	Glu	Asp	Val	Trp										
		675														

FIG. 5C

CATATGCGCT AATCTTCTTT TTCTTTTAT CACAGGAGAA ACTATCCCAC CCCCCTTCG	60
AAACACAATG ACAACTCCTG CGTAACTTGC AAATTCTTGT CTGACTAATT GAAAACTCCG	120
GACGAGTCAG ACCTCCAGTC AAACGGACAG ACAGACAAAC ACTTGGTGCG ATGTTCATAC	180
CTACAGACAT GTCAACGGGT GTTAGACGAC GGTTCCTTGC AAAGACAGGT GTTGGCATCT	240
CGTACGATGG CAACTGCGAG AGGTGTCGAC TTCTCCTTTA GGCAATAGAA AAAGACTAAG	300
AGAACAGCGT TTTTACAGGT TGCATTGGTT AATGTAGTAT TTTTITAGTC CCAGCATCTT	360
GTGGGTGCT CTGGGTTTCT AGAATAGGAA ATCACAGGAG AATGCAAATT CAGATGGAAG	420
AACAAAGAGA TAAAAACAA AAAAAACTG AGTTTTGCAC CAATAGAATG TTTGATGATA	480
TCATCCACTC GCTAAACGAA TCATGTGGT GATCTTCTCT TTAGTTTTGG TCTATCATAA	540
AACACATGAA AGTGAAATCC AAATACACTA CACTCCGGGT ATTGTCCCTC GTTTTACAGA	600
TGTCTCATTG TCTTACTTTT GAGGTCATAG GAGTTGCCTG TGAGAGATCA CAGAGATTAT	660
CACACTCACA TTTATCGTAG TTTCTATCT CATGCTGTGT GTCTCTGGTT GGTTCATGAG	720
TTTGGATTGT TGTACATTAA AGGAATCGCT GGAAAGCAAA GCTAACTAAA TTTCTTTGT	780
CACAGGTACG CTAACCTGTA AAACCTCACT GCCACGCCAG TCTTTCCTGA TTGGGCAAGT	840
GCACAACTA CAACCTGCAA AACAGCACTC CGCTTGTCAC AGGTGTCTC CTCTCAACCA	900
ACAAAAAAT AAGATTAAAC TTTCTTTGCT CATGCATCAA TCGGAGTTAT CTCTGAAAGA	960
GTTCCTTTG TGTAAATGTGT GCCAACTCA AACTGC AAA CTAACCACAG AATGATTTCC	1020
CTCACAAATTA TATAAATCA CCCACATTT CACAGACCGT AATTTCATGT CTCACTTTCT	1080
CTTTTGCTCT TCTTTTACTT AGTCAGTTTT GATAACTTCC TTTTITATTA CCTATCTTA	1140
TTTATTTATT TATTCATTTA TACCAACCAA CCAACCATGG CCACACAAGA AATCATCGAT	1200
TCTGTACTTC CGTACTTGAC CAAATGGTAC ACTGTGATTA CTGCAGCAGT ATTAGTCTTC	1260
CTTATCTCCA CAAACATCAA GAACTACGTC AAGGCAAAGA AATTGAAATG TGTGGATCCA	1320
CCATACTTGA AGGATGCCGG TCTCACTGGT ATTCTGTCTT TGATCGCCGC CATCAAGGCC	1380
AAGAACGACG GTAGATTGGC TAACTTTGCC GATGAAGTTT TCGACGAGTA CCCAAACCAC	1440
ACCTTCTACT TGTCTGTTGC CGGTGCTTTG AAGATTGTCA TGACTGTTGA CCCAGAAAAC	1500
ATCAAGGCTG TCTTGGCCAC CCAATTCCT GACTTCTCCT TGGGTACCAG ACACGCCCAC	1560

FIG. 6A

ES 2 298 276 T3

TTTGCTCCTT TGTTGGGTGA CGGTATCTTC ACCTTGGACG GAGAAGGTTG GAAGCACTCC	1620
AGAGCTATGT TGAGACCACA GTTTGCTAGA GACCAGATTG GACACGTTAA AGCCTTGGA	1680
CCACACATCC AAATCATGGC TAAGCAGATC AAGTTGAACC AGGGAAAGAC TTTCGATATC	1740
CAAGAATTGT TCTTTAGATT TACCGTCGAC ACCGCTACTG AGTTCCTGTT TGGTGAATCC	1800
GTTCACTCCT TGTACGATGA AAAATTGGGC ATCCCAACTC CAAACGAAAT CCCAGGAAGA	1860
GAAACTTTG CCGCTGCTTT CAACGTTTCC CAACACTACT TGGCCACCAG AAGTTACTCC	1920
CAGACTTTTT ACTTTTTGAC CAACCCTAAG GAATTCAGAG ACTGTAACGC CAAGGTCCAC	1980
CACTTGGCCA AGTACTTTGT CAACAAGGCC TTGAACITTA CTCCTGAAGA ACTCGAAGAG	2040
AAATCCAAGT CCGGTTACGT TTTCTTGATC GAATTGGTTA AGCAAACCAG AGATCCAAAG	2100
GTCTTGCAAG ATCAATTGTT GAACATTATG GTTGCCGGAA GAGACACCAC TGCCGGTTTG	2160
TTGTCCTTG CTTTGTTTGA ATTGGCTAGA CACCCAGAGA TGTGTTCCAA GTTGAGAGAA	2220
GAAATCGAAG TTAACITTTG TGTGGTGAA GACTCCCGCG TTGAAGAAAT TACETTCGAA	2280
GCCTGAAGA GATGTGAATA CTTGAAGGCT ATCCTTAACG AAACCTTGCG TATGTACCCA	2340
TCTGTTCTG TCAACTTTAG AACC GCCACC AGAGACACCA CTTGCCAAG AGGTGGTGGT	2400
GCTAACGGTA CCGACQCAAT CTACATTCCT AAAGGCTCCA CTGTTGCTTA CGTTGTCTAC	2460
AAGACCCACC GTTTGGAAGA ATACTACGGT AAGGACGCTA ACGACTTCAG ACCAGAAAGA	2520
TGGTTTGAAC CATCTACTAA GAAGTTGGGC TGGGCTTATG TTCCATTCAA CGGTGGTCCA	2580
AGAGTCTGCT TGGGTCAACA ATTCGCCTTG ACTGAAGCTT CTTATGTGAT CACTAGATTG	2640
GCCCAGATGT TTGAACTGT CTCATCTGAT CCAGGTCTCG AATACCCTCC ACCAAAGTGT	2700
ATTCACITGA CCATGAGTCA CAACGATGGT GTCTTTGTCA AGATGTAAAG TAGTCGATGC	2760
TGGGTATTGG ATTACATGTG TATAGGAAGA TTTTGGTTTT TTATTCGTTT TTTTTTTTAA	2820
TTTTTGTTAA ATTAGTTTAG AGATTTCAAT AATACATAGA TGGGTGCTAT TTCCGAAACT	2880
TTACTTCTAT CCCCTGTATC CTTATTATC CCTCTCAGTC ACATGATTGC TGTAAITGTC	2940
GTGCAGGACA CAACTCCCT AACGGACTTA AACCATAAAC AAGCTCAGAA CCATAAGCCG	3000
ACATCACTCC TTCTTCTCTC TTCTCCAACC AATAGCATGG ACAGACCCAC CCTCCTATCC	3060
GAATCGAAGA CCTTATTGA CTCCATACCC ACCTGGAAGC CCCTCAAGCC ACACACGTCA	3120
TCCAGCCAC CCATCACCAC ATCCCTCTAC TCGACAACGT CCAAAGACGG CGAGTTCTGG	3180

FIG. 6B

ES 2 298 276 T3

TGTGCCCGGA AATCAGCCAT CCCGGCCACA TACAAGCAGC CGTTGATTGC GTGCATACTC	3240
GGCGAGCCCA CAATGGGAGC CACGCATTGG GACCATGAAG CAAAGTACAT TCACGAGATC	3300
ACGGGTGTTT CAGTGTGCA GATTGAGAAG TTCGACGATG GATGGAAGTA CGATCTCGTT	3360
ECGGATTACG ACTTCGGTGG GTTGTTATCT AAACGAAGAT TCTATGAGAC GCAGCATGTG	3420
TTTCGGTTCG AGGATTGTGC GTACGTCATG AGTGTGCCTT TTGATGGACC CAAGGAGGAA	3480
GGTTACGTGG TTGGGACGTA CAGATCCATT GAAAGGTTGA GCTGGGGTAA AGACGGGGAC	3540
GTGGAGTGGG CCATGGCGAC GACGTCGGAT CCTGGTGGGT TTATCCCGCA ATGGATAACT	3600
CGATTGAGCA TCCCTGGAGC AATCGCAAAA GATGTGCCTA GTGTATTAAA CTACATACAG	3660
AAATAAAAAC GTGTCTTGAT TCATTGGTTT GGTTCCTGTT GGGTTCGAG CCAATATTTT	3720
ACATCATCTC CTAAATTCTC CAAGAATCCC AACGTAGCGT AGTCCAGCAC GCCCTCTGAG	3780
ATCTTATTTA ATATCGACTT CTCAACCACC GGTGGAATCC CGTTCAGACC ATTGTTACCT	3840
GTAGTGTGTT TGCTCTTGTT CTTGATGACA ATGATGTATT TGTCACGATA CCTGAAATAA	3900
TAAAACATCC AGTCATTGAG CTTATTACTC GTGAACTTAT GAAAGAACTC ATTCAAGCCG	3960
TTCCCAAAA ACCCAGAATT GAAGATCTTG CTCAACTGGT CATGCAAGTA GTAGATCGCC	4020
ATGATCTGAT ACTTTACCAA GCTATCCTCT CCAAGTTCTC CCACGTACGG CAAGTACGGC	4080
AACGAGCTCT GGAAGCTTTG TTGTTTGGGG TCATA	4115

FIG. 6C

```

gacctgtgac gcttccgggtg tcttgccacc agtctcttaag ttgaccgacg cccaagtcac 60
gtaccacttt atttccgggt acacttccaa gatgggtggg actgaagaag gtgtcacgga 120
accacaagct actttctccg ccctgttccg tcaaccattc ttgggtgttg acccaatgaa 180
gtacgctcaa caattgtctg acaagatctc gcaacacaag gctaacgcct ggttgttgaa 240
cacgggttgg gttggttctt ctgctgctag aggtggttaag agatgtctat tgaagtacac 300
cagagccatt ttggacgcta tccactctgg tgaatgtctc aagggtgaa acgaaacttt 360
cccagtcttc aacttgaatg tcccaacctc ctgtccagggt gtcccaagtg aaatcttgaa 420
cccaaccaag gcctggaccg gaagggtgtg actccttcaa caaggaaatc aagtctttgg 480
ctggtaagtt tgcctgaaaac tccaagaact atgttgacca agctaccgct gaagtgaag 540
ctgcagggtcc agaagcttaa agatatttat tcattattta gtctgcctat ttatttctca 600
ttacccatca tcattcaaca ctatatataa agttacttcg gatattcatt taatcgtgcg 660
tgtcgcaatt ggatgatttg gaactgcgct tgaaacggat tcatgcacga agcggagata 720
aaagattacg taatttatct cctgagacaa ttttagcgtt gtccacacgc ccttctttgt 780
tctgagcgaa ggataaataa ttgacttcc acagctcatt ctaatttccg tcacgcgaat 840
attgaagggg ggtacatgtg gccgctgaat gtgggggag taaacgcagt ctctctctc 900
ccaggaatag tgcaacggag gaaggataac ggatagaaag cggaaatgcga ggaaaatttt 960
gaacgcgcaa gaaaagcaat atccgggcta ccaggttttg agccaggga cacactccta 1020
tttctgtca atgactgaac atagaaaada caccaagacg caatgaaacg cacatggaca 1080
tttagacctc cccacatgtg atagtgtgtc ttaacagaaa agtataataa gaacccatgc 1140
cgtccctttt ctttgcgcgc tccaactttt ttttttttat cttacacaca tcacgacct 1200
gactgtacac gatattatcg ccacatactt caccaaatgg tacgtgatag taccactcgc 1260
tttgattgct tatagagtcc tgcactactt ctatggcaga tacttgatgt acaagcttgg 1320
tgctaaacca tttttccaga aacagacaga cggctgtttc ggattcaaag ctccgcttga 1380
attgttgaag aagaagagcg acggtaccct catagacttc acaotccagc gtatccacga 1440
tctcgatcgt cccgatatcc caactttcac attcccggtc ttttccatca accttgtcaa 1500
tacccttgag cgggagaaca tcaaggccat cttggccact cagtccaacg atttctctct 1560
gggtaccaga cactcgcact ttgctccttt gttgggtgat ggtatcttta cgttggatgg 1620
cgccggctgg aagcacagca gatctatgtt gagactacag tttgccagay aacagatttc 1680
ccacgtcaag ttgttggagc cacacgttca ggtgttcttc aaacacgtca gaaaggcaca 1740
gggcaagact tttgacatcc aggaattgtt tttcagattg accgtcgact ccgccaccga 1800
gtttttgttt ggtgaatccg ttgagtcctt gagagatgaa tctatcgssa tgtccatcaa 1860
tgcgcttgac tttgacggca aggttggbtt tgcctgatgt ttttaactatt cgcagaatta 1920
tttggtttcg agagcggtta tgcaacaatt gtaactgggt ttgaacggga aaaagttaa 1980
ggagtgaac gctaaagtgc acaagtttgc tgactactac gtcaacaagg ctttggactt 2040
gacgcctgaa caattggaaa agcaggatgg ttatgtgttt ttgtacgaat tgggtcaagca 2100
aaccagagac aagcaagtg: tgagagacca attgtgaac atcatggttg ctggtagaga 2160
caccaccgcc ggtttgttg: cgtttgtttt ctttgaattg gccagaaac cagaagttac 2220
caacaagttg agagaagaaa ttgaggaaca gtttggactc gg:agaaatg ctagtgttga 2280
agacatttcc tttgagtcgt tgaagtccct tgaatacttg aaggtgttc tcaacgaac 2340
cttgagattg taacctccg: tgccacagaa tttcagagt: gccatcaaga acactatcct 2400
cccaagaggt ggtggttaag: acgggttgtc tccg:tttg gtgagaaag: gtcagacct 2460
tatttacggt gtctacgca: cccacagaaa cccgctgtt taaggtaag: acgctctga 2520
gttttagacca gagagatgg: ttgagccaga gacaaagaa: cttggctgg: ccttctctcc 2580
attcaacggt ggtccaagaa tctgttttgg acagcagttt gccttgacag aagcttcgta 2640
tgtcactgtc aggttgcctc agtagtttgc acattgtct atggaccag acaccgaata 2700
tccacctaa: aaaatgtcgt: atttgacct gtct:tttc gatgggtgca atattgagat 2760
gtatttagagg gtcattgtt: attttgattg tttag:ttgt aattactga: taggttaatt 2820
catggattgt tatttattga taggggtttg cgtgtgtgc attcacttg gatcgttcca 2880
ggttgatgtt tctttcata: ctgtcgagtc aaaaggagtt ttg:tttg: aatccggagc 2940

```

FIG. 7A

ES 2 298 276 T3

gacctgtgac gcttcgggtg tcttgccacc agtctccaag ttgaccgacg cccaagtcac 60
 gtaccacttt atttccgggt acacttccaa gatggctggg actgaagaag gtgtcacgga 120
 accacaagct actttctccg ctigtcttcg tcaaccattc ttgggtgttg acccaatgaa 180
 gtacgctcaa caattgtctg acaagatctc gcaacacaag gctaacgcct gggtgttgaa 240
 caccgggttg gttgggtctt ctgctgctag aggtggtaag agatgctcat tgaagtacac 300
 cagagccatt ttggacgcta tccactctgg tgaattgtcc aaggttgaat acgaaacttt 360
 cccagctctc aacttgaatg tcccaacctc ctgtccaggt gtcccaagtg aaatcttgaa 420
 cccaaccaag gcttggaacg gaagggtgtg actccttcaa caaggaaac aagtctttgg 480
 ctggttaagt tgctgaaaac ttcaagacct atgtgacca agctaccgct gaagtgaag 540
 ctgcaggtcc agaagcttaa agatatttat tcattattta gtttgctat ttatttctca 600
 ttacccatca tcattcaaca ctatatataa agttacttcg gatattctg taatcgtgcg 660
 tgtcgcaatt ggatgatttg gaactgcgct tgaacggat tcatgcacga agcggagata 720
 aaagattacg taatttatct cctgagacaa ttttagccgt gttcacacgc cctctttgt 780
 tctgagcgaa ggataaataa ttagacttcc acagctcatt ctaatttccg tcacgcgaat 840
 attgaagggg ggtacatgtg gccgctgaat gtgggggcag taaacgcagt ctctctctc 900
 ccaggaatag tgcaacggag gaaggataac ggatagaaag cggaatgcga ggaaaatttt 960
 gaacgcgcaa gaaaagcaat atccgggcta ccaggttttg agccaggga cactctcta 1020
 tttctgctca atgactgaac atagaaaaaa caccaagacg caatgaaacg cacatggaca 1080
 ttttagacctc cccacatgtg atagtttgc ttaacagaaa agtataataa gaacccatgc 1140
 cgtccctttt ctttcgccgc ttcaactttt tttttttat cttacacaca tcacgacct 1200
 gactgtacac gatattatcg ccacatactt caccaatgg tacgtgatag taccactcgc 1260
 tttgattgct tatagagtcc tcgactactt ctatggcaga tacttgatgt acaagcttg 1320
 tgctaaacca ttttccaga aacagacaga cggctgtttc ggattcaaag ctccgcttga 1380
 attgttgaag aagaagagcg acggtaacct catagacttc aactccagc gtatccacga 1440
 tctcgatcgt cccgatatcc caactttcac attcccggtc tttccatca acctgtcaa 1500
 tacccttgag ccggagaaca tcaaggccat cttggccact cagtccaacg atttctctt 1560
 gggtaaccaga cactcgact ttgctcctt gttgggtgat ggtatcttta cgttgatgg 1620
 cgccggctgg aagcacagca gatctatgtt gagaccacag ttgcccagag aacagatttc 1680

ES 2 298 276 T3

```

atgtttttaa tagaagggtcg atctccatgt gattgttttg actgttactg tgattatgta 3000
atctgcggac gttatacaag catgtgattg tgggttttgca gccttttgca cgacaaatga 3060
tcgtcagacg attacgtaat ctttggttaga ggggtaaaaa aaaacaaaat ggcagccaga 3120
atttcaaaca ttctgcaaac aatgcaaaaa atgggaaact ccaacagaca aaaaaaaaaa 3180
ctccgcagca ctccgaaccc acagaacaat ggggcgccag aattattgac tattgtgact 3240
tttttacgct aacgctcatt gcagtgtagt gcgtcttaca cggggtattg ctttctacaa 3300
tgcaagggca cagttgaagg ttgcaacctt acgttgcccc gtgtcaactc aatttgacga 3360
gtaacttcct aagctcgaat tatgcagctc gtgcgtcaac ctatgtgcag gaaagaaaaa 3420
atccaaaaaa atcgaaaatg cgactttcga ttttgaaata accaaaaaga aaaatgtcgc 3480
acttttttct cgctctcgct ctctcgaccc aaatcacaaac aaatcctcgc gcgcagtatt 3540
tcgacgaaac cacaacaaat aaaaaaaaaa aattctacac cacttctttt tcttcaccag 3600
tcaacaaaaa acaacaaatt atacaccatt tcaacgattt ttgctcttat aaatgotata 3660
taatggttta attcaactca ggtatgttta ttttactggt ttcagctcaa gtatgttcaa 3720
atactaacta cttttgatgt ttgtcgctpt tctagaatca aaacaacgcc cacaacacgc 3780
cgagcttgtc gaatagacgg ttgttttact cattagatgg tcccagatta cttttcaagc 3840
caaagtctct cgagttttgt ttgtgttttc cccaattcct aactatgaag ggtttttata 3900
aggtccaaag accccaaggc atagtttttt tggttccttc ttgtcgtg 3948

```

FIG. 7B

GCTCAACAAT TGTCTGACAA GATCTCGCAA CACAAGGCTA ACGCCTGGTT GTTGAACACT	60
GGTTGGGTTG GTTCTTCTGC TGCTAGAGGT GGTAAGAGAT GTTCATTGAA GTACACCAGA	120
GCCATTTTGG ACGCTATCCA CTCTGGTGAA TTGTCCAAGG TTGAATACGA GACTTTCCCA	180
GTCTTCAACT TGAATGTCCC AACCTCCTGC CCAGGTGTCC CAAGTGAAAT CTTGAACCCA	240
ACCAAGGCCT GGACCGAAGG TGTTGACTCC TTCAACAAGG AAATCAAGTC TTTGGCTGGT	300
AAGTTTGCTG AAAACTTCAA GACCTATGCT GACCAAGCTA CCGCTGAAGT TAGAGCTGCA	360
GGTCCAGAAG CTTAAAGATA TTTATTCACT ATTTAGTTTG CCTATTATT TCTCATCACC	420
CATCATCATT CAACAATATA TATAAAGTTA TTTCGGAACT CATATATCAT TGTAAATCGT	480
CGTGTGCAA TTGGGTAAAT TGAAACTGTA GTTGGAAACGG ATTCATGCAC GATGCGGAGA	540
TAACACGAGA TTATCTCCTA AGACAATTTT GGCCTCATTC ACACGCCCTT CTTCTGAGCT	600
AAGGATAAAT AATTAGACTT CACAAGTTCA TTAAAATATC CGTCACGCGA AAAGTGCAAC	660
AATAAGGAAG GGGGGGGTAG ACGTAGCCGA TGAATGTGGG GTGCCAGTAA ACGCAGTCTC	720
TCTCTCCCCC CCCCCCCCCC CCCCCTCAGG AATAGTACAA CGGGGGAAGG ATAACGATA	780
GCAAGTGGA TCGGAGGAAA ATTTTGAATG CGCAAGGAAA GCAATATCCG GGCTATCAGG	840
TTTGTAGCCA GGGGACACAC TCCTCTTCTG CACAAAAACT TAACGTAGAC AAAAAAAAAA	900
AACTCCACCA AGACACAATG AATCGCACAT GGACATTTAG ACCTCCCCAC ATGTGAAAGC	960
TTCTCTGGCG AAAGCAAAA AAGTATAATA AGGACCCATG CCTTCCCTCT TCCTGGGCCG	1020
TTTCAACTTT TTCTTTTCT TGTCTATCA ACACACACAC ACCTCACGAC CATGACTGCA	1080
CAGGATATTA TCGCCACATA CATCACCAAA TGGTACGTGA TAGTACCACT CGCTTTGATT	1140
GCTTATAGGG TCCTCGACTA CTTTACGGC AGATACTTGA TGTACAAGCT TGGTGCTAAA	1200
CCGTTTTTCC AGAAACAAAC AGACGGTTAT TTCGGATTCA AAGCTCCACT TGAATTGTTA	1260
AAAAAGAAGA GTGACGGTAC CCTCATAGAC TTCACTCTCG AGCGTATCCA AGCGCTCAAT	1320
CGTCCAGATA TCCCAACTTT TACATTCCCA ATCTTTTCCA TCAACCTTAT CAGCACCTT	1380
GAGCCGGAGA ACATCAAGGC TATCTTGGCC ACCCAGTTCA ACGATTTCTC CTTGGGCACC	1440
AGACACTCGC ACTTTGCTCC TTTGTTGGGC GATGGTATCT TTACCTTGGA CGGTGCCGGC	1500
TGGAAGCACA GCAGATCTAT GTTGAGACCA CAGTTTGCCA GAGAACAGAT TTCCACGTC	1560

FIG. 8A

ES 2 298 276 T3

AAGTTGTTGG AGCCACACAT GCAGGTGTTT TCAAGCAGC TCAGAAAGGC ACAGGGCAAG	1620
ACTTTTGACA TCCAAGAATT GTTTTTCAGA TTGACCGTCG ACTCCGCCAC TGAGTTTTTG	1680
TTTGGTGAAT CCGTTGAGTC CTTGAGAGAT GAATCTATTG GGATGTCCAT CAATGCACCT	1740
GACTTTGACG GCAAGGCTGG CTTTGCTGAT GCTTTTAACT ACTCGCAGAA CTATTTGGCT	1800
TCGAGAGCCG TTATGCAACA ATTGTACTGG GTGTTGAACG GGAAAAAGTT TAAGGAGTGC	1860
AACGCTAAAG TGCACAAGTT TGCTGACTAT TACGTCAGCA AGGCTTTGGA CTTGACACCT	1920
GAACAATTGG AAAAGCAGGA TGGTTATGTG TTCTTGTACG AGTTGGTCAA GCAAACCAGA	1980
GACAGGCAAG TGTGAGAGA CCAGTTGTTG AACATCATGG TTGCCGGTAG AGACACCACC	2040
GCCGGTTTGT TGTGTTTTGT TTTCTTTGAA TTGGCCAGAA ACCCAGAGGT GACCAACAAG	2100
TTGAGAGAAG AAATCGAGGA CAAGTTTGGT CTTGGTGAGA ATGCTCGTGT TGAAGACATT	2160
TCCTTTGAGT CGTTGAAGTC ATGTGAATAC TTGAAGGCTG TTCTCAACGA AACTTTGAGA	2220
TTGTACCCAT CCGTGCCACA GAATTTTCTG GTTGCCACCA AAAACACTAC CCTTCCAAGG	2280
GGAGGTGGTA AGGACGGGTT ATCTCCTGTT TTGGTCAGAA AGGGTCAAAC CGTTATGTAC	2340
GGTGTCTACG CTGCCCACAG AAACCCAGCT GTCTACGGTA AGGACGCCCT TGAGTTTAGA	2400
CCAGAGAGGT GGTTTGAGCC AGAGACAAAG AAGCTTGGCT GGGCCTTCCT TCCATTCAAC	2460
GGTGGTCCAA GAATTTGCTT GGGACAGCAG TTGCGCTTGA CAGAAGCTTC GTATGTCACT	2520
GTCAGATTGC TCCAAGAGTT TGGACACTTG TCTATGGACC CCAACACCGA ATATCCACCT	2580
AGGAAAATGT CGCATTTGAC CATGTCCCTT TCGACGGTG CCAACATTGA GATGTATTAG	2640
AGGATCATGT GTTATTTTTG ATTGGTTTAG TCTGTTTGTG GCTATTGATT AGGTTAATTC	2700
ACGGATTGTT ATTTATTGAT AGGGGGTGCG TGTGTGTGTG TGTGTTGCAT TCACATGGGA	2760
TCGTTCCAGG TTGTTGTTTC CTCCATCCT GTTGAGTCAA AAGGAGTTTT GTTTTGTAAC	2820
TCCGGACGAT GTCTTAGATA GAAGGTCGAT CTCCATGTGA TTGTTTGACT GCTACTCTGA	2880
TTATGTAATC TGTAAGCCT AGACGTTATG CAAGCATGTG ATTGTGGTTT TTGCAACCTG	2940
TTTGACGAC AAATGATCGA CAGTCGATTA CGTAATCCAT ATTATTTAGA GGGGTAATAA	3000
AAATAAATG GCAGCCAGAA TTTCAAACAT TTTGCAAACA ATGCAAAAGA TGAGAACTC	3060
CAACAGAAAA AATAAAAAAA CTCCGACGCA CTCCGAACCA ACAAACAAT GGGGGGCGCC	3120
AGAATTATTG ACTATTGTGA CTTTTTTTAA TTTTTCCTGT TAACTTTCAT TGCAGTGAAG	3180

FIG. 8B

ES 2 298 276 T3

TGTGTTACAC	GGGGTGGTGA	TGGTGTGGT	TTCTACAATG	CAAGGGCACA	GTTGAAGGTT	3240
TCCACATAAC	GTTGCACCAT	ATCAACTCAA	TTTATCCTCA	TTCATGTGAT	AAAAGAAGAG	3300
CCAAAAGGTA	ATTGGCAGAC	CCCCCAAGGG	GAACACGGAG	TAGAAAGCAA	TGGAACACG	3360
CCCATGACAG	TGCCATTTAG	CCCAGAACAC	ATCTAGTATT	CTTTTTTTTT	TTTGTGCGCA	3420
GGTGCACACC	TGGACTTTAG	TTATTGCCCC	ATAAAGTTAA	CAATCTCACC	TTTGGCTCTC	3480
CCAGTGTCTC	CGCCTCCAGA	TGCTCGTTTT	ACACCCTCGA	GCTAACGACA	ACACAACACC	3540
CATGAGGGGA	ATGGGCAAAG	TTAAACACTT	TTGGTTTCAA	TGATTCCTAT	TTGCTACTCT	3600
CTTGTTTTGT	GTTTTGATTT	GCACCATGTG	AAATAAACGA	CAATTATATA	TACCTTTTCG	3660
TCTGTCTCC	AATGTCTCTT	TTTGCTGCCA	TTTTGCTTTT	TGCTTTTTCG	TTTTCACCTC	3720
TCTCCCACTC	CCACAATCAG	TGCAGCAACA	CACAA			3755

FIG. 8C

ES 2 298 276 T3

GACATCATAA TGACCCGGTT ATTTGCGCCT CAGGTTGCTT ATTTGAGCCG TAAAGTGCAG	60
TAGAAACTTT GCCTTGGGTT CAAACTCTAG TATAATGGTG ATAACGGTT GCACTCTTGC	120
CATAGGCATG AAAATAGGCC GTTATAGTAC TATATTTAAT AAGCGTAGGA GTATAGGATG	180
CATATGACCG GTTTTTCTAT ATTTTAAAGA TAATCTCTAG TAAATTTTGT ATTCTCAGTA	240
GGATTTCATC AAATTTTCGA ACCAATTCTG GCGAAAAAAT GATTCTTTTA CGTCAAAAGC	300
TGAATAGTGC AGTTTAAAGC ACCTAAAATC ACATATACAG CCTCTAGATA CGACAGAGAA	360
GCTCTTTATG ATCTGAAGAA GCATTAGAAT AGCTACTATG AGCCACTATT GGTGTATATA	420
TTAGGGATTG GTGCAATTAA GTACGTACTA ATAAACAGAA GAAATACTT AACCAATTTC	480
TGGTGTATAC TTAGTGGTGA GGGACCTTTT CTGAACATTC GGGTCAAAC TTTTTTTGGA	540
GTGCGACATC GATTTTTCGT TTGTGTAATA ATAGTGAACC TTTGTGTAAT AAATCTTCAT	600
GCAAGACTTG CATAATTCGA GCTTGGGAGT TCACGCCAAT TTGACCTCGT TCATGTGATA	660
AAAGAAAAGC CAAAAGGTAA TTAGCAGACG CAATGGGAAC ATGGAGTGA AAGCAATGGA	720
AGCACGCCCC GGACGGAGTA ATTTAGTCCA CACTACATCT GGGGGTTTTT TTTTGTGCG	780
CAAGTACACA CCTGGACTTT AGTTTTTGCC CCATAAAGTT AACAATCTAA CCTTTGGCTC	840
TCCAACCTCT TCCGCCCCCA AATATTCGTT TTTACACCCT CAAGCTAGCG ACAGCACAAC	900
ACCCATTAGA GGAATGGGGC AAAGTTAAAC ACTTTTGGCT TCAATGATTC CTATTCGCTA	960
CTACATTCTT CTCTTGTTTT GTGCTTTGAA TTGCACCATG TGAAATAAAC GACAATTATA	1020
TATACCTTTT CATCCCTCCT CCTATATCTC TTTTGTCTAC ATTTTGTITT TTACGTTTCT	1080
TGCTTTTGCA CTCTCCCACT CCCACAAAGA AAAAAAACT ACACTATGTC GTCTTCTCCA	1140
TCGTTTGCCC AAGAGGTTCT CGCTACCACT AGTCCTTACA TCGAGTACTT TCTTGACAAC	1200
TACACCAGAT GGTACTACTT CATACTTTG GTGCTTCTT CGTTGAACTT TATAAGTTG	1260
CTCCACACAA GGTACTTGGA ACGCAGGTT CACGCCAAGC CACTCGGTAA CTTTGTGAGG	1320
GACCTACGT TTGGTATCGC TACTCCGTTG CTTTGTATCT ACTTGAAGTC GAAAGGTACG	1380
GTCATGAAGT TTGCTTGGGG CCTCTGGAAC AACAAGTACA TCGTCAGAGA CCCAAAGTAC	1440
AAGACAAC TGCTCAGGAT TGTTGGCCTC CCATTGATTG AAACCATGGA CCCAGAGAAC	1500
ATCAAGGCTG TTTTGGCTAC TCAGTTCAAT GATTCTCTT TGGGAACCAG ACACGATTTT	1560

FIG. 9A

ES 2 298 276 T3

TTGTACTCCT TGTGCGGTGA CGGTATTTTC ACCTTGACG GTGCTGGCTG GAAACATAGT	1620
AGAACTATGT TGAGACCACA GTTTGCTAGA GAACAGGTTT CTCACGTCAA GTTGTGAG	1680
CCACACGTTT AGGTGTTCTT CAAGCACGTT AGAAAGCACC GCGGTAAAC GTTCGACATC	1740
CAAGAATTGT TCTTCAGGTT GACCGTCGAC TCCGCCACCG AGTTCTTGTT TGGTGAGTCT	1800
GCTGAATCCT TGAGGGACGA ATCTATTGGA TTGACCCCAA CCACCAAGGA TTTGATGGC	1860
AGAAGAGATT TCGCTGACGC TTTCAACTAT TCGCAGACTT ACCAGGCCTA CAGATTTTTG	1920
TTGCAACAAA TGTACTGGAT CTTGAATGGC TCGGAATTCA GAAAGTCGAT TGCTGTCTGT	1980
CACAAGTTTG CTGACCACTA TGTGCAAAAG GCTTTGGAGT TGACCGACGA TGACTTGCAG	2040
AAACAAGACG GCTATGTGTT CTTGTACGAG TTGGCTAAGC AAACCAGAGA CCCAAAGGTC	2100
TTGAGAGACC AGTTATTGAA CATTTTGGTT GCCGGTAGAG ACACGACCGC CGGTTTGTG	2160
TCATTTGTTT TCTACGAGTT GTCAAGAAAC CCTGAGGTGT TTGCTAAGTT GAGAGAGGAG	2220
GTGGAACACA GATTTGGACT CGGTGAAGAA GCTCGTGTG AAGAGATCTC GTTTGAGTCC	2280
TTGAAGTCTT GTGAGTACTT GAAGGCTGTC ATCAATGAAA CCTTGAGATT GTACCCATCG	2340
GTTCACACA ACTTTAGAGT TGCTACCAGA AACACTACCC TCCCAAGAGG TGGTGGTGAA	2400
GATGGATACT CGCCAATTGT CGTCAAGAAG GGTCAAGTTG TCATGTACAC TGTTATTGCT	2460
ACCCACAGAG ACCCAAGTAT CTACGGTGCC GACGCTGACG TCTTCAGACC AGAAAGATGG	2520
TTTGAACCAG AAAC TAGAAA GTTGGGCTGG GCATACGTTT CATTCAATGG TGGTCCAAGA	2580
ATCTGTTTGG GTCAACAGTT TGCCTTGACC GAAGCTTCAT ACGTCACTGT CAGATTGCTC	2640
CAGGAGTTTG CACACTTGTC TATGGACCCA GACACCGAAT ATCCACCAA ATTGCAGAAC	2700
ACCTTGACCT TGTGCTCTT TGATGGTGCT GATGTTAGAA TGTACTAAGG TTGCTTTTCC	2760
TTGCTAATTT TCTTCTGTAT AGCTTGTGTA TTAAATTGA ATCGGCAATT GATTTTCTG	2820
ATACCAATAA CCGTAGTGCG ATTTGACCAA AACCGTTCAA ACTTTTGTG CTCTCGTTGA	2880
CGTGCTCGCT CATCAGCACT GTTTGAAGAC GAAAGAGAAA ATTTTTTGTA AACAACTG	2940
TCCAAATTA CCCAACGTGA ACCATTATGC AAATGAGCGG CCCTTCAAC TGGTCGCTGG	3000
AAGCATTCGG GGATATCTAC AACGCCCTTA AGTTTGAAAC AGACATTGAT TTAGACACCA	3060
TAGATTTTCA CGGCATCAAG AATGACCTTG CCCACATTTT GACGACCCCA ACACCACTGG	3120
AAGAATCACG CCAGAACTA GGCGATGGAT CCAAGCCTGT GACCTTGCCC AATGGAGACG	3180

FIG. 9B

ES 2 298 276 T3

AAGTGGAGTT GAACCAAGCG TTCCTAGAAG TTACCACATT ATTGTGGAAT GAGTTTGACT	3240
TGGACCAATT GAACGCGGCA GAGTTGTTAT ACTACGCTGG CGACATATCC TACAAGAAGG	3300
GCACATCAAT CGCAGACAGT GCCAGATTGT CTTATTATTT GAGAGCAAAC TACATCTTGA	3360
ACATACTTGG GTATTTGATT TCGAAGCAGC GATTGGATTT GATAGTCACG GACAACGACG	3420
CGTTGTTTGA TAGTATTTTG AAAAGTTTGG AAAAGATCTA CAAGTTGATA AGCGTGTTGA	3480
ACGATATGAT TGACAAGCAA AAGGTGACAA GCGACATCAA CAGTCTAGCA TTCATCAATT	3540
GCATCAACTA CTCGAGAGGT CAACTATTCT CCGCACACGA ACTTTTGGGA CTGGTTTTGT	3600
TTGGATTGGT CGACATCTAT TTCAACCAGT TTGGCACATT AGACAACTAC AAGAAGGTAT	3660
TGGCATTGAT ACTGAAGAAC ATCAGCGATG AAGACATCTT GATCATAAC TCCCTCCCAT	3720
CGACACTACA ATTGTTTAAG CTGGTGTTGG ACAAGAAAGA CGACGCTGCA GTTGAACAGT	3780
TCTACAAGTA CATCACTTCA ACAGTGTCAC GAGACTACAA CTCCAACATC GGCTCCACAG	3840
CCAAAGATGA TATCGATTGG TCCAAAACCA AACTCAGTGG CTTTGAGGTG TTGACGAGTT	3900

FIG. 9C

ES 2 298 276 T3

CCTGCAGAAT TCGCGGCCGC GTCGACAGAG TAGCAGTTAT GCAAGCATGT GATTGTGGTT	50
TTTGCAACCT GTTTGCACGA CAAATGATCG ACAGTCGATT ACGTAATCCA TATTATTTAG	120
AGGGGTAATA AAAAATAAAT GGCAGCCAGA ATTTCAAACA TTTTGCAAAC AATGCAAAAG	180
ATGAGAAACT CCAACAGAAA AAATAAAAAA ACTCCGCAGC ACTCCGAACC AACAAAACAA	240
TGGGGGGCGC CAGAATTATT GACTATTGTG ACTTTTTTTT ATTTTTTCCG TTAACTTTCA	300
TTGCAGTGAA GTGTGTTACA CGGGGTGGTG ATGGTGTGGG TTTCTACAAT GCAAGGGCAC	360
AGTTGAAGGT TTCCACATAA CGTTGCACCA TATCAACTCA ATTTATCCTC ATTCATGTGA	420
TAAAGAAGA GCCAAAAGGT AATTGGCAGA CCCCCAAGG GGAACACGGA GTAGAAAGCA	480
ATGGAAACAC GCCCATGACA GTGCCATTTA GCCCACAACA CATCTAGTAT TCTTTTTTTT	540
TTTTGTGCGC AGGTGCACAC CTGGACTTTA GTTATTGCCC CATAAAGTTA ACAATCTC	600
CTTTGGCTCT CCCAGTGTCT CCGCCTCCAG ATGCTCGTTT TACACCTCG AGCTAACGAC	660
AACACAACAC CCATGAGGGG AATGGGCAA GTTAAACACT TTTGGTTTCA ATGATTCCTA	720
TTTGCTACTC TCTTGTTTTG TGTTTTGATT TGCACCATGT GAAATAAACG ACAATTATAT	780
ATACCTTTTC GTCTGTCTCT CAATGTCTCT TTTTGTGCTC ATTTTGCTTT TTGCTTTTTG	840
CTTTTGCACT CTCTCCCACT CCCACAATCA GTGCAGCAAC ACACAAAGAA GAAAAATAAA	900
AAAACCTACA CTATGTCGTC TTCTCCATCG TTTGCTCAGG AGGTTCTCGC TACCACTAGT	960
CCTTACATCG AGTACTTTCT TGACAACTAC ACCAGATGGT ACTACTTCAT CCCTTTGGTG	1020
CTTCTTTCGT TGAAC TTCAT CAGCTTGCTC CACACAAAGT ACTTGGAAACG CAGGTTCCAC	1080
GCCAAGCCGC TCGGTAACGT CGTGTGGAT CCTACGTTTG GTATCGCTAC TCCGTTGATC	1140
TTGATCTACT TAAAGTCGAA AGGTACAGTC ATGAAGTTTG CCTGGAGCTT CTGGAACAAC	1200
AAGTACATTG TCAAAGACCC AAAGTACAAG ACCACTGGCC TTAGAATTGT CGGCCTCCCA	1260
TTGATTGAAA CCATAGACCC AGAGAACATC AAAGCTGTGT TGGCTACTCA GTTCAACGAT	1320
TTCTCCTTGG GAACTAGACA CGATTTCTTG TACTCCTTGT TGGGCGATGG TATTTTTTACC	1380
TTGGACGGTG CTGGCTGGAA ACACAGTAGA ACTATGTTGA GACCACAGTT TGCTAGAGAA	1440
CAGGTTTCCC ACGTCAAGTT GTTGGAACCA CACGTTCAAG TGTTCCTCAA GCACGTTAGA	1500
AAACACCGCG GTCAGACTTT TGACATCCAA GAATTGTCTC TCAGATTGAC CGTCGACTCC	1560

FIG. 10A

ES 2 298 276 T3

GCCACCGAGT TCTTGTTTGG TGAGTCTGCT GAATCCTTGA GAGACGACTC TGTTGGTTTG	1620
ACCCCAACCA CCAAGGATTT CGAAGGCAGA GGAGATTTCG CTGACGCTTT CACTACTCG	1680
CAGACTTACC AGGCCTACAG ATTTTGTG TG CAACAAATGT ACTGGATTTT GAATGGCGCG	1740
GAATTCAGAA AGTCGATTGC CATCGTGAC AAGTTTGCTG ACCACTATGT GCAAAAGGCT	1800
TTGGAGTTGA CCGACGATGA CTTGCAGAAA CAAGACGGCT ATGTGTTCTT GTACGAGTTG	1860
GCTAAGCAAA CTAGAGACCC AAAGGTCTTG AGAGACCACT TGTGGAACAT TTTGGTTGCC	1920
GGTAGAGACA CGACCGCCGG TTTGTTGTCG TTTGTGTTCT ACGAGTTGTC GAGAAACCTT	1980
GAAGTGTTTG CCAAGTTGAG AGAGGAGGTG GAAAACAGAT TTGGACTCGG CGAAGAGGCT	2040
CGTGTTGAAG AGATCTCTTT TGAGTCCTTG AAGTCCTGTG AGTACTTGAA GGCTGTCATC	2100
AATGAAGCCT TGAGATTGTA CCCATCTGTT CCACACAACT TCAGAGTTGC CACCAGAAAC	2160
ACTACCCTTC CAAGAGGCGG TGGTAAAGAC GGATGCTCGC CAATTGTTGT CAAGAAGGGT	2220
CAAGTTGTCA TGTACACTGT CATTGGTACC CACAGAGACC CAAGTATCTA CGGTGCCGAC	2280
GCCGACGTCT TCAGACCAGA AAGATGGTTC GAGCCAGAAA CTAGAAAGTT GGGCTGGGCA	2340
TATGTTCCAT TCAATGGTGG TCCAAGAATC TGTTTGGGTC AGCAGTTTGC CTTGACTGAA	2400
GCTTCATACG TCACTGTCAG ATTGCTCCAA GAGTTTGGAA ACTTGTCCTT GGATCCAAAC	2460
GCTGAGTACC CACCAAAATT GCAGAACACC TTGACCTTGT CACTCTTTGA TGGTGTGAC	2520
GTTAGAATGT TCTAAGGTTG CTTATCCTTG CTAGTGTTAT TTATAGTTTG TGTATTTAAA	2580
TTGAATCGGC GATTGATTTT TCTGGTACTA ATAACGTAG TGGGTTTTGA CCAAAACCGT	2640
TCAAACTTTT TTTTTTTTTT TCTTCCCCCT ACCTTCGTTG CTCGCTCATC AGCACTGTTT	2700
GAAAACGAAA AAAGAAAATT TTTGTAAAC AACATTGCCC AACTTACCC AACGTGAACC	2760
ATTATAACCA AATGAGCGGC GCTTCAACT GGTCACTGGA GGCATTCGGG GATATCTACA	2820
ACACCCTTAA GTTTGAGGAA GACATTGATT TAGACACCAT AGATTTGAGC GGCATCAAGA	2880
ATGACCTTGT CCACATTTTG ACAACCCCAA CACCACTGGA AGAATCGCGC CAGAACTAG	2940
GCGATGGATC CAAGCCTGTG GCCTTGCCCA ATGGAGACGA AGTGGAGTTG AACCAAGCGT	3000
TCCTAGAAGT TACCACATTA TTGTGGAACG AGTTTGACTT GGACCAATTG AACGCGGCCG	3060
AGTTGTTATA CTACGCCGGC GACATATCCT ACAAGAAGGG CACATCAATT GCCGACAGTG	3120
CCAGATTGTC TTAATATTTG AGAGCAAACT ACATCTTGAA CATACTTGGG TACTTTATTT	3180

FIG. 10B

ES 2 298 276 T3

CGAAGCAGCG ATTGGATGTG ATAGTCACCG ACAACAACGC GTTGTTTGAT AATATTTTGA	3240
AAAGTTTGA AAAGATCTAC AAGTTGATAA GCGCGTTGAA CGATATGATT GACAAGCAAA	3300
AGGTGACAAG CGACATCAAC AGTCTAGCAT TTATCAACTG CATCAACTAC TCGAGGGGTC	3360
AACTATTCTC CGCACACGAA CTTTGTGGAC TGGTTTGTGTT TGGATTGGTT GACAACTATT	3420
TCAACCAGTT TGGCTCATT GACAACTACA AGAAAGTATT GGCATTGATA CTGAAGAACA	3480
TCAGTGATGA AGATATCTTG ATCGTACGCT TCCTCCCATC GACACTACAA TTGTTTAAGC	3540
TGGTGTGGA TAAGAAAGAC GACGCCACTG TTGACCAGTT CTACAAGTAC ATCACCTCAA	3600
CAGTGTGCA AGACTACAAC TCCAACATCG GAGCCACAGC CAAAGATGAT ATCGATTGT	3660
CCAAAGCC	3668

FIG. 10C

ES 2 298 276 T3

GATGTGGTGC TTGATTTCTC GAGACAÇATC CTTGTGAGGT GCCATGAATC TGTACCTGTC	60
TGTAAGCACA GGGAACTGCT TCAACACCTT ATTGCATATT CTGTCTATTG CAAGCGTGTG	120
CTGCAACGAT ATCTGCCAAG GTATATAGCA GAACGTGCTG ATGGTTCCTC CGGTCAATATT	180
CTGTTGGTAG TTCTGCAGGT AAAATTGGAT GTCAGGTAGT GGAGGGAGGT TTGTATCGGT	240
TGTGTTTTCT TCTTCCTCTC TCTCTGATTC AACCTCCACG TCTCCTTCGG GTTCTGTGTC	300
TGTGTCTGAG TCGTACTGTT GGATTAAGTC CATCGCATGT GTGAAAAAA GTAGCGCTTA	360
TTTAGACAAC CAGTTCGTTG GCGGGGTATC AGAAATAGTC TGTTGTGCAC GACCATGAGT	420
ATGCAACTTG ACGAGACGTC GTTAGGAATC CACAGAATGA TAGCAGGAAG CTTACTACGT	480
GAGAGATTCT GCTTAGAGGA TGTTCCTCTC TTGTTGATTC CATTAGGTGG GTATCATCTC	540
CGGTGGTGAC AACTTGACAC AAGCAGTTCC GAGAACCACC CACAACAATC ACCATTCCAG	600
CTATCACTTC TACATGTCAA CCTACGATGT ATCTCATCAC CATCTAGTTT CTTGGCAATC	660
GTTTATTTGT TATGGGTCAA CATCCAATAC AACTCCACCA ATGAAGAAGA AAAACGGAAA	720
GCAGAATACC AGAATGACAG TGTGAGTTCC TGACCATTGC TAATCTATGG CTATATCTAG	780
TTTGCTATCG TGGGATGTGA TCTGTGTCGT CTTCAATTCG GTTTGTGTTT ATTTCCGGTA	840
TGAATATTGT TATACTAAAT ACTTGATGCA CAAACATGGC GCTCGAGAAA TCGAGAATGT	900
GATCAACGAT GGGTTCTTTG GGTTCCGCTT ACCTTTGCTA CTCATGCGAG CCAGCAATGA	960
GGGCCGACTT ATCGAGTTCA GTGTCAAGAG ATTCGAGTCG GCGCCACATC CACAGAACAA	1020
GACATTGGTC AACCGGGCAT TGAGCGTTCC TGTGATACTC ACCAAGGACC CAGTGAATAT	1080
CAAAGCGATG CTATCGACCC AGTTTGATGA CTTTTCCTT GGGTTGAGAC TACACCAATT	1140
TGCGCCGTTG TTGGGGAAAG GCATCTTTAC TTTGGACGGC CCAGAGTGGG AGCAGAGCCG	1200
ATCTATGTTG CGTCCGCAAT TTGCCAAAGA TCGGGTTTCT CATATCCTGG ATCTAGAACC	1260
GCATTTTGTG TTGCTTCGGA AGCACATTGA TGGCCACAAT GGAGACTACT TCGACATCCA	1320
GGAGCTCTAC TTCCGGTTCT CGATGGATGT GCGGACGGGG TTTTGTGTTG GCGAGTCTGT	1380
GGGGTCGTTG AAAGACGAAG ATGCGAGGTT CCTGGAAGCA TTCAATGAGT CGCAGAAGTA	1440
TTTGGCAACT AGGGCAACGT TGCACGAGTT GTACTTTCTT TGTGACGGGT TTAGGTTTCG	1500
CCAGTACAAC AAGGTTGTGC GAAAGTTCTG CAGCCAGTGT GTCCACAAGG CGTTAGATGT	1560

FIG. 11A

ES 2 298 276 T3

TGCACCGGAA	GACACCAGCG	AGTACGTGTT	TCTCCGCGAG	TTGGTCAAAC	ACACTCGAGA	1620
TCCCGTTGTT	TTACAAGACC	AAGCGTTGAA	CCTCTTGCTT	GCTGGACGCG	ACACCACCGC	1680
GTCGTTATTA	TCGTTTGCAA	CATTTGAGCT	AGCCCGGAAT	GACCACATGT	GGAGGAAGCT	1740
ACGAGAGGAG	GTTATCCTGA	CGATGGGACC	GTCCAGTGAT	GAAATAACCG	TGGCCGGGTT	1800
GAAGAGTTGC	CGTTACCTCA	AAGCAATCCT	AAACGAAACT	CTTCGACTAT	ACCCAAGTGT	1860
GCCTAGGAAC	GCGAGATTTG	CTACGAGGAA	TACGACGCTT	CCTCGTGGCG	GAGGTCCAGA	1920
TGGATCGTTT	CCGATTTTGA	TAAGAAAGGG	CCAGCCAGTG	GGGTATTTCA	TTTGTGCTAC	1980
ACACTTGAAT	GAGAAGGTAT	ATGGGAATGA	TAGCCATGTG	TTTCGACCGG	AGAGATGGGC	2040
TGCGTTAGAG	GGCAAGAGTT	TGGGCTGGTC	GTATCTTCCA	TTCAACGGCG	GCCCGAGAAG	2100
CTGCCTTGGT	CAGCAGTTTG	CAATCCTTGA	AGCTTCGTAT	GTTTTGGCTC	GATTGACACA	2160
GTGCTACACG	ACGATACAGC	TTAGAACTAC	CGAGTACCCA	CCAAAGAAAC	TCGTTTCATCT	2220
CACGATGAGT	CTTCTCAACG	GGGTGTACAT	CCGAAGTAGA	ACTTGATTAT	GTGTTTATGG	2280
TTAATCGGGG	CAAAGCACTG	CAAGTCATTG	ATGTTTGTGG	AAGCCCAGCA	TTGGTGTTCG	2340
GGAGCATCAA	TAACCAATGT	CTTGAAGGGT	TTGATTTTCT	TGACCTTCTT	CTTCCTGAGC	2400
TTCTTTCCGT	CAAACCTGTA	CAGAATGGCC	ATCATTTCAG	GAACAACCAC	GTACGACGGC	2460
CGGTACCGCA	TCTGGAGTAT	CTCGCCGTGG	TTCAAGTAGC	ACGAAAACAG	CAACGACGTC	2520
ACCATCTGCT	TCCCAATCTT	GACACCCACA	GATACCCCTG	CGGCTTCATG	GATCAAAAAC	2580
GTGCGCAACC	CCGCGTATAT	GTCCATGTAA	TTCTCCATGG	CCACCTCCAT	CAACACACTG	2640
ATGGAGCGAC	TGACGGTGCC	ACCACTGCCC	TCGGTTGAGT	CAAGGCAGTA	TGATGCCGGG	2700
ATCCAGTACT	CCAATGGGAA	CCTCTGCACG	GTGTGCTGTC	AGTTTTTGAG	GCGTATTTCC	2760
ATCCATGATC	GTTCTTTGGT	GCTGTAGTAT	AACGAGCTCT	TGGTGTCTCT	GAAATGGAAC	2820
AGGTTGGATG	TGTTGTTGAG	TTTGTCTGCG	TGCTTGGTTT	GCAAGTCTTC	GATCGAGCGT	2880
AGTGAGTAGA	CAGTTGGCGG	GGGTGGTGCC	TCGGGCTTTA	TTCTGTGTTT	GTGTTTCCTT	2940
CTTAGTCTTG	GAATGACGCT	GTTATCGACG	GTTCGTAGTA	TAAGTAGCGC	CAATATGAGA	3000
ATGTATATCC	GCATCAGCCA	AGACTCTTCA	GCCTGTTACA	ACGACTGAGG	CTGTTGGCCG	3060
TGTGACCAAT	TGGTTTCTTT	GGTGACCTAG	ATTGGTCCCG	CAGGGAAAGC	AAGGGCTGCT	3120
AGGGGGGCAT	ACCAAACAAG	GTGCTGTAAT	CAGTATCTAT	GGTGCTACCA	TGTGTGTGGT	3180

FIG. 11B

ES 2 298 276 T3

TGGGGGAAA TTCCCGCATT TTGTGTAAAC GAAAGTTCTA GAAAGTTCTC GTGGGTTCTG	3240
AGAATCTGCT GGAACCATCC ACCCGCATT CCGTTGCCAA AGTGGAAGA GCAATCAACC	3300
CACCCTGCTT TGCCCAATCA GCCATTCCCC TGGGAATATA AATTCAAC	3348

FIG. 11C

ES 2 298 276 T3

TGGAGTCGCC	AGACTTGCTC	ACTTTTGA	CTCTTCGAAA	CTCAAAGTAC	GTTTCAGGCGG	60
TGCTCAACGA	AACGCTCCGT	ATCTACCCGG	GGGTACCACG	AAACATGAAG	ACAGCTACGT	120
GCAACACGAC	GTTGCCACGC	GGAGGAGGCA	AAGACGGCAA	GGAACCTATC	TTGGTGCAGA	180
AGGGACAGTC	CGTTGGGTTG	ATTACTATTG	CCACGCAGAC	GGACCCAGAG	TATTTTGGGG	240
CCGACGCTGG	TGAGTTTAAG	CCGGAGAGAT	GGTTTGATTG	AAGCATGAAG	AACTTGGGGT	300
GTAATACTT	GCCGTTCAAT	GCTGGGCCAC	GGACTTGCTT	GGGGCAGCAG	TACACTTTGA	360
TTGAAGCGAG	CTACTTGCTA	GTCCGGTTGG	CCCAGACCTA	CCGGGCAATA	GATTTGCAGC	420
CAGGATCGGC	GTACCCACCA	AGAAAGAAGT	CGTTGATCAA	CATGAGTGCT	GCCGACGGGG	480
TGTTTGTAAG	GCTTTATAAG	GATGTAACGG	TAGATGGATA	GTTGTGTAGG	AGGAGCGGAG	540
ATAAATTAGA	TTTGATTTTG	TGTAAGGTTT	TGGATGTCAA	CCTACTCCGC	ACTTCATGCA	600
GTGTGTGTGA	CACAAGGGTG	TACTACGTGT	GCCTGTGCGC	CAAGAGACAG	CCCAAGGGGG	660
TGGTAGTGTG	TGTTGGCGGA	AGTGCATGTG	ACACAACGCG	TGGTTCTTGG	CCAATGGTGG	720
ACTAAGTGCA	GGTAAGCAGC	GACCTGAAAC	ATTCTCTAAC	GCTTAAGACA	CTGGTGGTAG	780
AGATGCGGAC	CAGGCTATTG	TTGTCGTGCT	ACCCGGCGCA	TGGAAAATCA	ACTGCGGGAA	840
GAATAAATTT	ATCCGTAGAA	TCCACAGAGC	GGATAAATTT	GCCCACCTCC	ATCATCAACC	900
ACGCCGCCAC	TAACCTACATC	ACTCCCCTAT	TTTCTCTCTC	TCTCTTTGTC	TTACTCCGCT	960
CCCGTTTCCT	TAGCCACAGA	TACACACCCA	CTGCAACAG	CAGCAACAAT	TATAAAGATA	1020
CGCCAGGCCC	ACCTTCTTTT	TTTTTCTTCA	CTTTTTTGAC	TGCAACTTTT	TACAATCCAC	1080
CACAGCCACG	ACCACAGCCG	CTATGATTGA	ACAACCTCTA	GAATATTGGT	ATGTCGTTGT	1140
GCCAGTGTGG	TACATCATCA	AACAACCTCT	TGCATACACA	AAGACTCGCG	TCTTGATGAA	1200
AAAGTTGGGT	GCTGCTCCAG	TCACAAACAA	GTTGTACGAC	AACGCTTTTG	GTATCGTCAA	1260
TGGATGGAAG	GCTCTCCAGT	TCAAGAAAGA	GGGCAGGGCT	CAAGAGTACA	ACGATTACAA	1320
GTTTGACCAC	TCCAAGAACC	CAAGCGTGGG	CACCTACGTC	AGTATTCTTT	TCGGCACCAG	1380
GATCGTCGTG	ACCAAAGATC	CAGAGAATAT	CAAAGCTATT	TTGGCAACCC	AGTTTGCTGA	1440
TTTTTCTTTG	GGCAAGAGGC	ACACTCTTTT	TAAGCCTTTG	TTAGGTGATG	GGATCTTCAC	1500
ATTGGACGGC	GAAGGCTGGA	AGCACAGCAG	AGCCATGTTG	AGACCACAGT	TTGCCAGAGA	1560

FIG. 12A

ES 2 298 276 T3

ACAAGTTGCT CATGTGACGT CGTTGGAACC ACAC TTCAG TTGTTGAAGA AGCATATTCT	1620
TAAGCACAAG GGTGAATACT TTGATATCCA GGAATTGTTC TTTAGATTTA CCGTTGATTG	1680
GGCCACGGAG TTCTTATTTG GTGAGTCCGT GCACTCCTTA AAGGACGAAT CTATTGGTAT	1740
CAACCAAGAC GATATAGATT TTGCTGGTAG AAAGGACTTT GCTGAGTCGT TCAACAAAGC	1800
CCAGGAATAC TTGGCTATTA GAACCTTGGT GCAGACGTC TACTGGTTGG TCAACAACAA	1860
GGAGTTTAGA GACTGTACCA AGCTGGTGCA CAAGTTCACC AACTACTATG TTCAGAAAGC	1920
TTTGGATGCT AGCCAGAAG AGCTTGAAAA GCAAAGTGGG TATGTGTTCT TGTACGAGCT	1980
TGTCAAGCAG ACAAGAGACC CCAATGTGTT GCGTGACCAG TCTTTGAACA TCTTGTGGC	2040
CGGAAGAGAC ACCACTGCTG GGTGTGTGTC GTTGCTGTC TTTGAGTTGG CCAGACACCC	2100
AGAGATCTGG GCCAAGTTGA GAGAGGAAAT TGAACAACAG TTGGTCTTG GAGAGACTC	2160
TCGTGTGAA GAGATTACCT TTGAGAGCTT GAAGAGATGT GAGTACTTGA AAGCGTTCCT	2220
TAATGAAACC TTGCGTATTT ACCCAAGTGT CCCAAGAAAC TTCAGAATCG CCACCAAGAA	2280
CACGACATTG CCAAGGGGCG GTGGTTCAGA CGGTACCTCG CCAATCTTGA TCCAAAAGGG	2340
AGAAGCTGTG TCGTATGGTA TCAACTCTAC TCATTGGAC CCGTCTATT ACGGCCCTGA	2400
TGCTGCTGAG TTCAGACCAG AGAGATGGTT TGAGCCATCA ACCAAAAGC TCGGCTGGGC	2460
TTACTTGCCA TTCAACGGTG GTCCAAGAAT CTGTTGGGT CAGCAGTTTG CCTTGACGGA	2520
AGCTGGCTAT GTGTGGTTA GATTGGTGCA AGAGTCTCC CACGTTAGGC TGGACCCAGA	2580
CGAGGTGTAC CCGCCAAAGA GGTTGACCAA CTTGACCATG TGTTCAGG ATGGTGCTAT	2640
TGTCAAGTTT GACTAGCGGC GTGGTGAATG CGTTTGATT TGTAGTTTCT GTTTCAGTA	2700
ATGAGATAAC TATTCAGATA AGGCGAGTGG ATGTACGTTT TGTAAAGATT TCCTTACAAC	2760
CTTGGTGGGG TGTGTGAGGT TGAGGTTGCA TCTTGGGGAG ATTACACCTT TTGCAGCTCT	2820
CCGTATACAC TTGTACTCTT TGTAACTCT ATCAATCATG TGGGGGGGGG GGTTCATTGT	2880
TTGGCCATGG TGGTGCATGT TAAATCCGCC AACTACCCAA TCTCACATGA AACTCAAGCA	2940
CACTAAAAAA AAAAAAGATG TTGGGGGAAA ACTTTGGTTT CCCTTCTTAG TAATTAAACA	3000
CTCTCACTCT CACTCTCACT CTCTCCACTC AGACAAACCA ACCACCTGGG CTGCAGACAA	3060
CCAGAAAAAA AAGAACAAA ATCCAGATAG AAAAACAAAG GGCTGGACAA CCATAAATAA	3120
ACAATCTAGG GTCTACTCCA TCTTCCACTG TTTCTCTTC TTCAGACTTA GCTAACAAAC	3180

FIG. 12B

AACTCACTTC ACCATGGATT ACGCAGGCAT CACGCGTGGC TCCATCAGAG GCGAGGCCTT	3240
GAAGAAACTC GCAGAATTGA CCATCCAGAA CCAGCCATCC AGCTTGAAAG AAATCAACAC	3300
CGGCATCCAG AAGGACGACT TTGCCAAGTT GTTGTCTGCC ACCCCGAAAA TCCCCACCAA	3360
GCACAAGTTG AACGGCAACC ACGAATTGTC TGAGGTCGCC ATTGCCAAAA AGGAGTACGA	3420
GGTGTTCGATT GCCTTGAGCG ACGCCACAAA AGACCCAATC AAAGTGACCT' CCCAGATCAA	3480
GATCTTGATT GACAAGTTCA AGGTGTACTT GTTTGAGTTG CCTGACCAGA AGTTCTCCTA	3540
CTCCATCGTG TCCAACATCCG TCAACATCGC CCCCTGGACC TTGCTCGGGG AGAAGTTGAC	3600
CACGGGCTTG ATCAACTTGG CTTTCCAGAA CAACAAGCAG CACTTGGACG AGGTCATTGA	3660
CATCTTCAAC GAGTTCATCG ACAAGTTCTT TGGCAACACG GAGCCGCAAT TGACCAACTT	3720
CTTGACCTTG TCGGGTGTGT TGGACGGGTT GATTGACCAT GCCAACTTCT TGAGCGTGTC	3780
CTCGCGGACC TTCAAGATCT TCTTGAACCT GGACTCGTAT GTGGAC	3826

FIG. 12C

ES 2 298 276 T3

TTACAATCAT GGAGCTCGCT AGGAACCCAG ATGTCTGGGA GAAGCTCCGC GAAGAGGTCA	50
ACACGAACCTT TGGCATGGAG TCGCCAGACT TGCTCACTTT TGACTCTCTT AGAAGCTCAA	120
AGTACGTTCA GCGGGTGCTC AACGAAACGC TTCGTATCTA CCCGGGGGTG CCACGAAACA	180
TGAAGACAGC TACGTGCAAC ACGACGTTGC CGCGTGGAGG AGGCAAAGAC GGTAAGGAAC	240
CTATTTTGGT GCAGAAGGGC CAGTCCGTTG GGTGATTAC TATTGCCACG CAGACGGACC	300
CAGAGTATTT TGGGGCAGAT GCTGGTGAGT TCAAACCGGA GAGATGGTTT GATTCAAGCA	360
TGAAGAACTT GGGGTGTAAG TACTTGCCGT TCAATGCTGG GCCCCGGA CTGTTGGGGC	420
AGCAGTACAC TTTGATTGAA GCGAGCTATT TGCTAGTCAG GTTGGCGCAG ACCTACCGGG	480
TAATCGATTT GCTGCCAGGG TCGGCGTACC CACCAAGAAA GAAGTCGTTG ATCAATATGA	540
GTGCTGCCGA TGGGGTGGTT GTAAAGTTTC ACAAGGATCT AGATGGATAT GTAAGGTGTG	600
TAGGAGGAGC GGAGATAAAT TAGATTTGAT TTTGTGTAAG GTTTAGCACG TCAAGCTACT	660
CCGCACTTTG TGTGTAGGGA GCACATACTC CGTCTGCGCC TGTGCCAAGA GACGGCCCAG	720
GGGTAGTGTG TGGTGGTGGG AGTGCATGTG ACACAATACC CTGGTCTTGG CCAATTGGGG	780
ATTTAGTGTA GGTAAGCTGC GACCTGAAAC ACTCCTCAAC GCTTGAGACA CTGGTGGTA	840
GAGATGCGGG CCAGGAGGCT ATTCTTGTCTG TGCTACCCGT GCACGGAAAA TCGATTGAGG	900
GAAGAACAAA TTTATCCGTG AAATCCACAG AGCGGATAAA TTTGTCACAT TGCTGCGTTG	960
CCCACCCACA GCATTCTCTT TTCTCTCTCT TTGTCTTACT CCGCTCCTGT TTCCTTATCC	1020
AGAAATACAC ACCAACTCAT ATAAAGATAC GCTAGCCCAG CTGTCTTTCT TTTTCTTCA	1080
TTTTTTTGGT GTGTTGCTTT TTTGGCTGCT ACTTTCTACA ACCACCACCA CCACCACCAC	1140
CATGATTGAA CAAATCCTAG AATATTGGTA TATTGTTGTG CCTGTGTTGT ACATCATCAA	1200
ACAACTCATT GCCTACAGCA AGACTCGCGT CTTGATGAAA CAGTTGGGTG CTGCTCCAAT	1260
CACAAACCAG TTGTACGACA ACGTTTTCGG TATCGTCAAC GGATGGAAGG CTCTCCAGTT	1320
CAAGAAAGAG GGCAGAGCTC AAGAGTACAA CGATCACAAG TTTGACAGCT CCAAGAACCC	1380
AAGCGTCGGC ACCTATGTCA GTATTCTTTT TGGCACCAAG ATTGTCGTGA CCAAGGATCC	1440
AGAGAATATC AAAGCTATTT TGGCAACCCA GTTTGGCGAT TTTTCTTTGG GCAAGAGACA	1500
CGCTCTTTTT AAACCTTTGT TAGGTGATGG GATCTTCACC TTGGACGGCG AAGGCTGGAA	1560

FIG. 13A

ES 2 298 276 T3

GCATAGCAGA	TCCATGTTAA	GACCACAGTT	TGCCAGAGAA	CAAGTTGCTC	ATGTGACGTC	1620
GTTGGAACCA	CACTTCCAGT	TGTTGAAGAA	GCATATCCTT	AAACACAAGG	GTGAGTACTT	1680
TGATATCCAG	GAATTGTTCT	TTAGATTAC	TGTCGACTCG	GCCACGGAGT	TCTTATTTGG	1740
TGAGTCCGTG	CACTCCTTAA	AGGACGAAAC	TATCGGTATC	AACCAAGACG	ATATAGATTT	1800
TGCTGCTAGA	AAGGACTTTG	CTGAGTCGTT	CAACAAAGCC	CAGGAGTATT	TGTCTATTAG	1860
AATTTTGGTG	CAGACCTTCT	ACTGGTTGAT	CAACAACAAG	GAGTTTAGAG	ACTGTACCAA	1920
GCTGGTGAC	AAGTTTACCA	ACTACTATGT	TCAGAAAGCT	TTGGATGCTA	CCCCAGAGGA	1980
ACTTGAAAAG	CAAGGCGGGT	ATGTGTTCTT	GTATGAGCTT	GTCAAGCAGA	CGAGAGACCC	2040
CAAGGTGTTG	CGTGACCAGT	CTTGAACAT	CTTGTTGGCA	GGAAGAGACA	CCACTGCTGG	2100
GTTGTTGTCC	TTTGCTGTGT	TTGAGTTGGC	CAGAAACCCA	CACATCTGGG	CCAAGTTGAG	2160
AGAGGAAATT	GAACAGCAGT	TTGGTCTTGG	AGAAGACTCT	CGTGTTGAAG	AGATTACCTT	2220
TGAGAGCTTG	AAGAGATGTG	AGTACTTGAA	AGCGTTCCTT	AACGAAACCT	TGCGTGTTTA	2280
CCCAAGTGTC	CCAAGAAACT	TCAGAATCGC	CACCAAGAAT	ACAACATTGC	CAAGGGGTGG	2340
TGGTCCAGAC	GGTACCCAGC	CAATCTTGAT	CCAAAGGGGA	GAAGGTGTGT	CGTATGGTAT	2400
CAACTCTACC	CACTTAGATC	CTGTCTATTA	TGGCCCTGAT	GCTGCTGAGT	TCAGACCAGA	2460
GAGATGGTTT	GAGCCATCAA	CCAGAAAGCT	CGGCTGGGCT	TACTTGCCAT	TCAACGGTGG	2520
GCCACGAATC	TGTTTGGGTC	AGCAGTTTGC	CTTGACCGAA	GCTGGTTACG	TTTGGTTCAG	2580
ATTGGTGCAA	GAGTTCTCCC	ACATTAGGCT	GGACCCAGAT	GAAGTGTATC	CACCAAAGAG	2640
GTTGACCAAC	TTGACCATGT	GTTTGCAGGA	TGGTGCTATT	GTCAAGTTTG	ACTAGTACGT	2700
ATGAGTGCGT	TTGATTTTGT	AGTTTCTGTT	TGCAGTAATG	AGATAACTAT	TCAGATAAGG	2760
CGGGTGGATG	TACGTTTTGT	AAGAGTTTCC	TTACAACCCT	GGTGGGTGTG	TGAGGTTGCA	2820
TCTTAGGGAG	AGATAGCACC	TTTTGCAGCT	CTCCGTATAC	AGTTTTACTC	TTTGTAACCT	2880
ATGCCAATCA	TGTGGGGATT	CATTGTTTGC	CCATGGTGGT	GCATGC AAAA	TCCCCCAAC	2940
TACCCAATCT	CACATGAAAC	TCAAGCACAC	TAGAAAAAAA	AGATGTTGCG	TGGGTTCTTT	3000
TGATGTTGGG	GAAAACTTTC	GTTTCCTTTC	TCAGTAATTA	AACGTTCTCA	CTCAGACAAA	3060
CCACCTGGGC	TGCAGACAAC	CAGAAAAAAC	AAAATCCAGA	TAGAAGAAGA	AAGGGCTGGA	3120
CAACCATAAA	TAAACAACCT	AGGGTCCACT	CCATCTTTCA	CTTCTTCTTC	TTCAGACTTA	3180

FIG. 13B

ES 2 298 276 T3

TCTAACAAAC	GACTCACTTC	ACCATGGATT	ACGCAGGTAT	CACGCGTGGG	TCCATCAGAG	3240
GCGAAGCCTT	GAAGAAACTC	GCCGAGTTGA	CCATCCAGAA	CCAGCCATCC	AGCTTGAAAG	3300
AAATCAACAC	CGGCATCCAG	AAGGACGACT	TTGCCAAGTT	GTTGTCTTCC	ACCCCGAAAA	3360
TCCACACCAA	GCACAAGTTG	AATGGCAACC	ACGAATTGTC	CGAAGTCGCC	ATTGCCAAAA	3420
AGGAGTACGA	GGTGTGATT	GCCTTGAGCG	ACGCCACGAA	AGAACCAATC	AAAGTCACCT	3480
CCCAGATCAA	GATCTTGATT	GACAAGTTCA	AGGTGTACTT	GTTTGAGTTG	CCCGACCAGA	3540
AGTTCTCCTA	CTCCATCGTG	TCCAACTCGG	TTAACATTGC	CCCCTGGACC	TTGCTCGGTG	3600
AGAAGTTGAC	CACGGGCTTG	ATCAACTTGG	CGTTCCAGAA	CAACAAGCAG	CACCTGGACG	3660
AAGTCATCGA	CATCTTCAAC	GAGTTCATCG	ACAAGTTCTT	TGGCAACACA	GAGCCGCAAT	3720
TGACCAACTT	CFTGACCTTG	TCCGGTGTGT	TGGACGGGTT	GATTGACCAT	GCCAACTTCT	3780
TGAGCGTGTC	CTCCAGGACC	TTCAAGATCT	TCTTGAACCT	GGACTCGTTT	GTGGACAACT	3840
CGGACTTCTT	GAACGACGTG	GAGAACTACT	CCGACTTTTT	GTACGACGAG	CCGAACGAGT	3900
ACCAGAACTT						3910

FIG. 13C

GAATTCCTTG	GATCTAATTC	CAGCTGATCT	TGCTAATCCT	TATCAACGTA	GTTGTGATCA	60
TTGTTTGTCT	GAATTATACA	CACCAAGTGA	AGAATATGGT	CTAATTGCA	CGTCCCCTG	120
GCATTGTGTG	TTTGTGGGGG	GGGGGGGGTG	CACACATTTT	TAGTGCCATT	CTTTGTTGAT	180
TACCCCTCCC	CCCTATCATT	CATTCCCACA	GGATTAGTTT	TTTCCTCACT	GGAATTCGCT	240
GTCCACCTGT	CAACCCCCCC	CCCCCCCCCC	CCCACTGCCC	TACCCTGCCC	TGCCCTGCAC	300
GTCCTGTGTT	TTGTGCTGTG	TCTTTCCQAC	GCTATAAAAG	CCCTGGCGTC	CGGCCAAGGT	360
TTTTCCACCC	AGCCAAAAAA	ACAGTCTAAA	AAATTTGGTT	GATCCTTTTT	GGTTGCAAGG	420
TTTTCCACCA	CCACTTCCAC	CACCTCAACT	ATTCGAACAA	AAGATGCTCG	ATCAGATCTT	480
ACATTACTGG	TACATTGTCT	TGCCATTGTT	GGCCATTATC	AACCAGATCG	TGGCTCATGT	540
CAGGACCAAT	TATTTGATGA	AGAAATTGGG	TGCTAAGCCA	TTACACACAG	TCCAACGTGA	600
CGGGTGGTTG	GGCTTCAAAT	TCGGCCGTGA	ATTCCTCAAA	GCAAAAAGTG	CTGGGAGACT	660
GGTTGATTTA	ATCATCTCCC	GTTTCCACGA	TAATGAGGAC	ACTTTCTCCA	GCTATGCTTT	720
TGGCAACCAT	GTGGTGTTCA	CCAGGGACCC	CGAGAATATC	AAGGCGCTTT	TGGCAACCCA	780
GTTTGGTGAT	TTTTCAATTG	GCAGCAGGGT	CAAGTTCTTC	AAACCATTAT	TGGGGTACGG	840
TATCTTCACA	TTGSAQGCCG	AAGGCTGGAA	GCACAGCAGA	GCCATGTTGA	GACCACAGTT	900
TGCCAGAGAA	CAAGTTGCTC	ATGTGACGTC	GTTGGAACCA	CACCTCCAGT	TGTTGAAGAA	960
GCATATCCTT	AAACACAAGG	GTGAGTACTT	TGATATCCAG	GAATTGTTCT	TTAGATTTAC	1020
TGTCGACTCG	GCCACGGAGT	TCTTATTTGG	TGAGTCCGTG	CACCTCCTAA	AGGACGAGGA	1080
AATTGGCTAC	GACACGAAAG	ACATGTCTGA	AGAAAGACGC	AGATTTGCCG	ACCGGTTCAA	1140
CAAGTCGCAA	GTCTACGTGG	CCACCAGAGT	TGCTTTACAG	AACTTGTA CT	GGTTGGTCAA	1200
CAACAAAGAG	TTCAAGGAGT	GCAATGACAT	TGTCCACAAG	TTTACCAACT	ACTATGTTCA	1260
GAAAGCCTTG	GATGCTACCC	CAGAGGAACT	TGAAAAGCAA	GGCGGGTATG	TGTTCTTGTA	1320
TGAGCTTGTC	AAGCAGACGA	GAGACCCCAA	GGTGTTCGCT	GACCAGTCTT	TGAACATCTT	1380
GTTGGCAGGA	AGAGACACCA	CTGCTGGGTT	GTTGTCTTTT	GCTGTGTTTG	AGTTGGCCAG	1440
AAACCCACAC	ATCTGGGCCA	AGTTGAGAGA	GGAAATTGAA	CAGCAGTTTG	GTCTTGAGAA	1500
AGACTCTCGT	GTTGAAGAGA	TTACCTTTGA	GAGCTTGAAG	AGATGTGAGT	ACTTGAAGGC	1560

FIG. 14A

ES 2 298 276 T3

CGTGTGAAC GAAACTTTGA GATTACACCC AAGTGTCCCA AGAAACGCAA GATTTGCGAT	1620
TAAAGACACG ACTTTACCAA GAGGCGGTGG CCCCACGGC AAGGATCCTA TCTTGATCAG	1680
GAAGGATGAG GTGGTGCAGT ACTCCATCTC GGCAACTCAG ACAAATCCTG CTTATTATGG	1740
CGCCGATGCT GCTGATTTTA GACCGGAAAG ATGGTTTGAA CCATCAACTA GAAACTTGGG	1800
ATGGGCTTTC TTGCCATTCA ACGGTGGTCC AAGAATCTGT TTGGGACAAC AGTTTGCTTT	1860
GACTGAAGCC GGTACGTTT TGGTTAGACT TGTTTCAGGAG TTTCCAAACT TGTCACAAGA	1920
CCCCGAAACC AAGTACCCAC CACCTAGATT GGCACACTTG ACGATGTGCT TGTTTGACGG	1980
TGCACACGTC AAGATGTCAT AGGTTTCCCC ATACAAGTAG TTCAGTAATT ATACACTGTT	2040
TTTACTTTCT CTTCATACCA AATGGACAAA AGTTTTAAGC ATGCCTAACA ACGTGACCGG	2100
ACAATTGTGT CGCACTAGTA TGTAACAATT GTAAAAATAG TGTACACTAA TTTGTGGTGG	2160
CCGGAGATAA ATTACAGTTT GGTTTTGTGT AAACCTCGCG ATATCTCTGG CAGTTTCTCT	2220
TCTCCGACG AGCTTTGCCA CGGGTTTGCT CTGGGGCCAA CAATTCAAA AGGGGGAGAA	2280
ACTTAACACC CCTTATCTCT CCACTCTAGG TTGTAGCTCT TGTGGGGATG CAATTGTCTG	2340
ACGTTTTTTA TGTTTTGTCT AGACTTTGAT GATTACGTTG GATTTCTTAT GTCTGAGGCG	2400
TGCTTGAAAG AAGTGTCAAA ATGTGACAGG CGACGCTATT GGACATGAAC GCGAAAGGGT	2460
TATTTGCATC AATACGAGGG GCTGACTCTA GTCTAGGATG GCAGTCCTAG GTTGCAAACA	2520
TGTTGCACCA TATCCCTCCT GGAGTTGGTC GACCTCGCCT ACGCCACCCT CAGCGATCGG	2580
CACTTTCCGT TGTTCAATAT TTCTCCTTCC CATTGTTCCA GGGGTTATCA ACAACGTTGC	2640
CGGCCTCCTC CCCAAATTAC AAGAAAAATA AATTGTGCA CGGCACCGAT CTGTCAAAGA	2700
TACAGATAAA CCTTAAATCT GCAAAAACAA GACCCCTCCC CATAGCCTAG AAGCACCAGC	2760
AAGATGATGG AGCAACTCCT CCACTACTGG TACATCGCAC TCTCTGTATG GTTCATCCTT	2820
CGCTACTTGG CTTCCACGC ACGAGCCGTC TACTTGCGCC ACAAGCTCGG CGCGGCGCCA	2880
TTCACGCACA CCCAGTACGA CGGCTGGTAT GGGTTCAAGT TTGGGCGGGA GTTTCTCAAG	2940
GCGAAGAAGA TCGGGCGGCA GACGGACTTG GTGCATGCGC GGTTCCTGG CGGCATGGAC	3000
ACCTTCTCGA GCTACACTTT CGGCATCCAT ATCATCCTTA CCCGGGACCC GGAGAACATC	3060
AAGGCGGTCT TGGCGACGA GTTCGATGAC TTCTCGCTCG GTGGCAGGAT CAGGTTCTTG	3120
AAGCCGTTGT TGGGGTATGG GATATTCACG	3150

FIG. 14B

ES 2 298 276 T3

AAAACCGATA CAAGAAGAAG ACAGTCAACA AGAACGTTAA TGTCAACCAG GCGCCAAGAA	60
GACGGTTTGG CGGACTTGGA AGAATGTGGC ATTGGOCCAT GATGTTTATG TTCTGGAGAG	120
GTTTTTCAAG GAATCGTCAT CCTCCGCCAC CACAAGAACC ACCAGTTAAC GAGATCCATA	180
TTCACAACC ACCGCAAGGT GACAATGCTC AACAAACAACA GCAACAACAA CAACCCCCAC	240
AAGAACAGTG GAATAATGCC AGTCAACAAA GAGTGGTGAC AGACGAGGGA GAAAACGCAA	300
GCAACAGTGG TTCTGATGCA AGATCAGCTA CACCGCTTCA TCAGGAAAAG CAGGAGCTCC	360
CACCACCATA TGCCCATCAC GAGCAACACC AGCAGGTTAG TGTATAGTAG TCTGTAGTTA	420
AGTCAATGCA ATGTACCAAT AAGACTATCC CTTCTTACAA CCAAGTTTTC TGCCGCGCCT	480
GTETGGCAAC AGATGCTGGC CGACACACTT TCAACTGAGT TTGGTCTAGA ATTCTTGAC	540
ATGCACGACA AGGAAACTCT TACAAAGACA ACACTTGTGC TCTGATGCCA CTGATCTTG	600
CTAAGCCTTA TCAACGTAAT TGAGATCATT GTTTGTCTGA ATTATACACA CCAGTGGGAG	660
AATCTGGTCT AATCTGCACG CCTCATGGGC ATTGTGTGTT TTGGGGGGGG GGGGGGGGGT	720
GCACACATTT TTAGTGCAGAA TGTTTGTTTG CTGGTTCCCC CTCCCCCTC CCCCCTATCA	780
TGCCACAGG ATTAGTTTTT TCCTCACTGG AATTGCTGT CCACCTGTCA ACCCCCTCAC	840
TGCCCTGCCC TGCCCTGCAC GCCCTGTGTT TTGTGCTGTG GCACTCCAC GCTATAAAG	900
CCCTGGCGTA CGGCCAAGGT TTTTCCTCAC AGCCAAAAA AATTGCGCT GATCCTTTTG	960
GGCTGCAAGG TTTTTCACCA CCACCACCAC CACCACCTCA ACTATTCAA CAAAGGATGC	1020
TCGACCAGAT CTTCCATTAC TGGTACATTG TCTTGCCATT GTTGGTCATT ATCAAGCAGA	1080
TCGTGGCTCA TGCCAGGACC AATTATTTGA TGAAGAAGTT GGGCGCTAAG CCATTACAC	1140
ATGTCCAAC AGACGGGTGG TTTGGCTTCA AATTGGCCG TGAATTCCTC AAAGCTAAAA	1200
GTGCTGGGAG GCAGGTTGAT TTAATCATCT CCCGTTTCCA CGATAATGAG GACACTTTCT	1260
CCAGCTATGC TTTTGGCAAC CATGTGGTGT TCACCAGGGA CCCCAGAAAT ATCAAGGCGC	1320
TTTTGGCAAC CCAGTTTGGT GATTTTTCAT TGGGAAGCAG GGTCAAATTC TTCAAACCAT	1380
TGTTGGGGTA CGGTATCTTC ACCTTGACG GCGAAGGCTG GAAGCACAGC AGAGCCATGT	1440
TGAGACCACA GTTTGCCAGA GAGCAAGTTG CTCATGTGAC GTCGTTGGAA CCACATTTCC	1500
AGTTGTTGAA GAAGCATATT CTTAAGCACA AGGGTGAATA CTTTGATATC CAGGAATTGT	1560

FIG. 15A

ES 2 298 276 T3

TCTTTAGATT TACCGTTGAT TCAGCGACGG AGTTCTTATT TGGTGAGTCC GTGCACTCCT	1620
TAAGGGACGA GGAAATTGGC TACGATACGA AGGACATGGC TGAAGAAAGA CGCAAATTTG	1680
CCGACGCGTT CAACAAGTCG CAACTCTATT TGTCCACCAG AGTTGCTTTA CAGACATTGT	1740
ACTGGTTGGT CAACAACAAA GAGTTC AAGG AGTGCAACGA CATTGTCCAC AAGTTCACCA	1800
ACTACTATGT TCAGAAAGCC TTGGATGCTA CCCCAGAGGA ACTTGAAAAA CAAGGCGGGT	1860
ATGTGTTCTT GTACGAGCTT GCCAAGCAGA CGAAAGACCC CAATGTGTTG CGTGACCAGT	1920
CTTTGAACAT CTGTGTTGGCT GGAAGGGACA CCACTGCTGG GTTGTGTTCC TTTGCTGTGT	1980
TTGAGTTGGC CAGGAACCCA CACATCTGGG CCAAGTTGAG AGAGGAAATT GAATCACACT	2040
TTGGGCTGGG TGAGGACTCT CGTGTGGAAG AGATTACCTT TGAGAGCTTG AAGAGATGTG	2100
AGTACTTGAA AGCGTGTGTT ACGAAACGT TGAGATTACA CCCAAGTGT CCAAGAAACG	2160
CAAGATTTGC GATTAAAGAC ACGACTTTAC CAAGAGGCGG TGGCCCCAAC GGC AAGGATC	2220
CTATCTTGAT CAGAAAGAAT GAGGTGGTGC AATACTCCAT CTCGGCAACT CAGACAAATC	2280
CTGCTTATTA TGGCGCCGAT GCTGCTGATT TTAGACCGGA AAGATGGTTT GAGCCATCAA	2340
CTAGAACTT GGGATGGGCT TACTTGCCAT TCAACGGTGG TCCAAGAATC TGCTTGGGAC	2400
AACAGTTTGC TTTGACCGAA GCCGGTTACG TTTTGGTTAG ACTTGTTT CAG GAATTCCTTA	2460
GCTTGTCACA GGACCCCGAA ACTGAGTACC CACCACCTAG ATTGGCACAC TTGACGATGT	2520
GCTTGTTTGA CGGGGCATAC GTCAAGATGC AATAGGTTTT GGTGTTGACTT TGTGTTCCATA	2580
TGCAAGTAGT TCAGTAATTA CACACTAATT TGTGGTGCC GCGGATAAAT TACCGTTTGG	2640
TTTTGTGTAA AAATTCGGAC ATCTCTGGTG GTTCCCTTC TCCGAGCAG CTTTGCCACG	2700
GGTTTGCTCT GCGGCCAACA AATTCGAAAG GGGGGGGGGG GGGGGAGAAA GTTAACACCC	2760
CCTGTTCCTA CCGTAGGCTG TAGCTCTTGT GGGGGGATGT AATTGTCGTA CGTTTTCATG	2820
TTTGGCCAG ACTTTGATGA TTACGTAGGC TTTCTTATGT CTAAGGCGTG CTTGACACAA	2880
GTGTCAAAAG GTGACAGGCG ACGTTATTCG ACATGAACGC AAAAGGGTAA TTTGCATCGA	2940
TACGAGGGGT TGCCTCTGGT CTAAGAAGGA CCCCCCAGGT TGCAAACATG TTGCACTGCA	3000
TCCCACTCAG AGTTGGTCGA CCACGCCTAC GCTTACCCTC AGCGATCGGC ACTTTCCGTT	3060
GCTCAATATT TCTCTCCCC CTGCTTCCCC CCATTGTTCC AGGGATTATC AACAAAGTTG	3120
CCGGTCTCCT CTCCCCCCC TCCCCCAGT TATGTACAAG AAAATTAAAT TGTCGCACGG	3180

FIG. 15B

ES 2 298 276 T3

CACCGATACG	TCAAAGATAC	AGAGAAACCT	TAATCCCTCC	CATAGCCTAG	AAGCATCAAA	3240
AAGATGATTG	AGCAACTCCT	CCAGTACTGG	TACATTGCAC	TCCCTGTATG	GTTCAATTCTC	3300
CGCTACGTGG	CTTCCCACGC	ACGAACCATC	TACTTGCGCC	ACAAGCTCGG	CGCGGCGCCG	3360
TTCACGCACA	CCCAGTACGA	CGGATGGTAT	GGGTTCAAGT	TTGGGCGGGA	GTTTCTCAAG	3420
GCGAAGAAGA	TTGGAAGGCA	GACGGACTTG	GTGCATGCGC	GGTTCCGTGG	AGGGGGCATG	3480
GATACTTTCT	CGAGCTATAC	TTTCGGCATC	CATATCAATC	TTACTCGGGA	CCCGGAGAAC	3540
ATCAAGGCGG	TCCTGGCGAC	GCAGTTCGAT	GACTTTTCG			3579

FIG. 15C

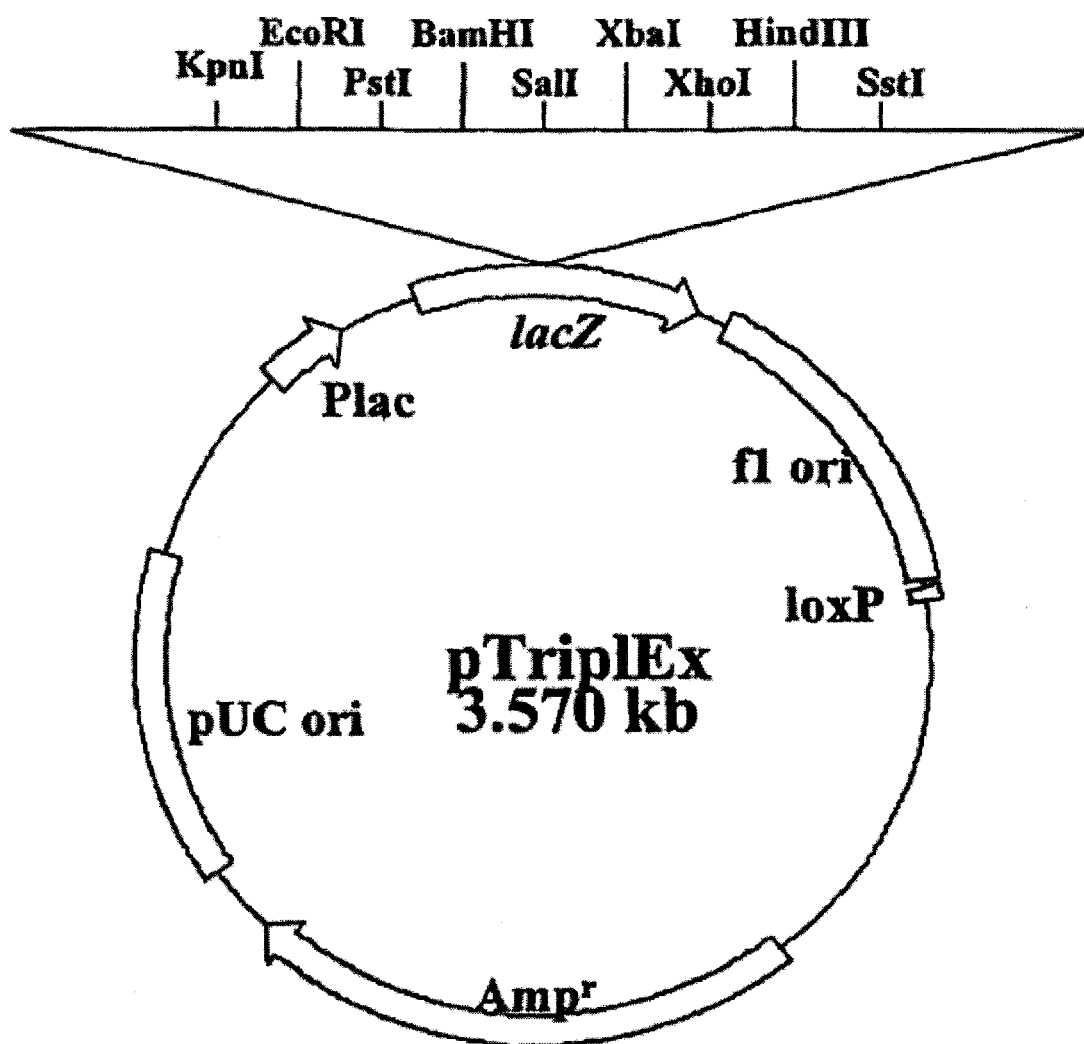


FIG. 16

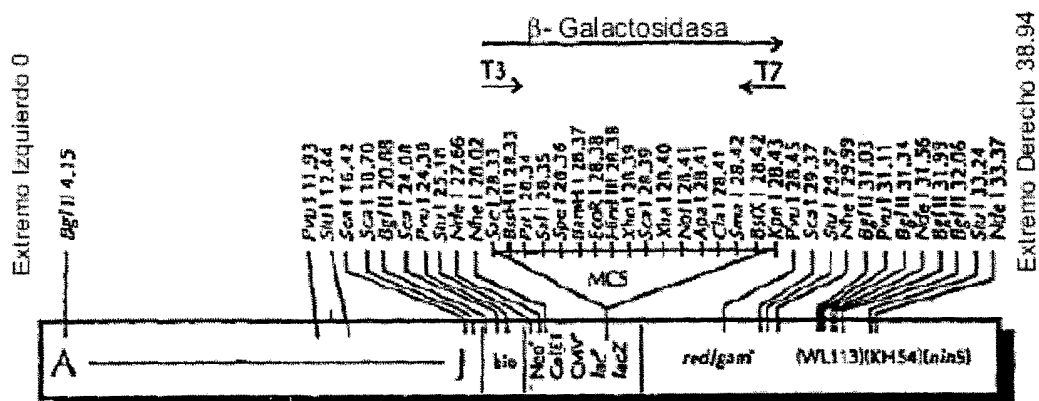


FIG. 17A

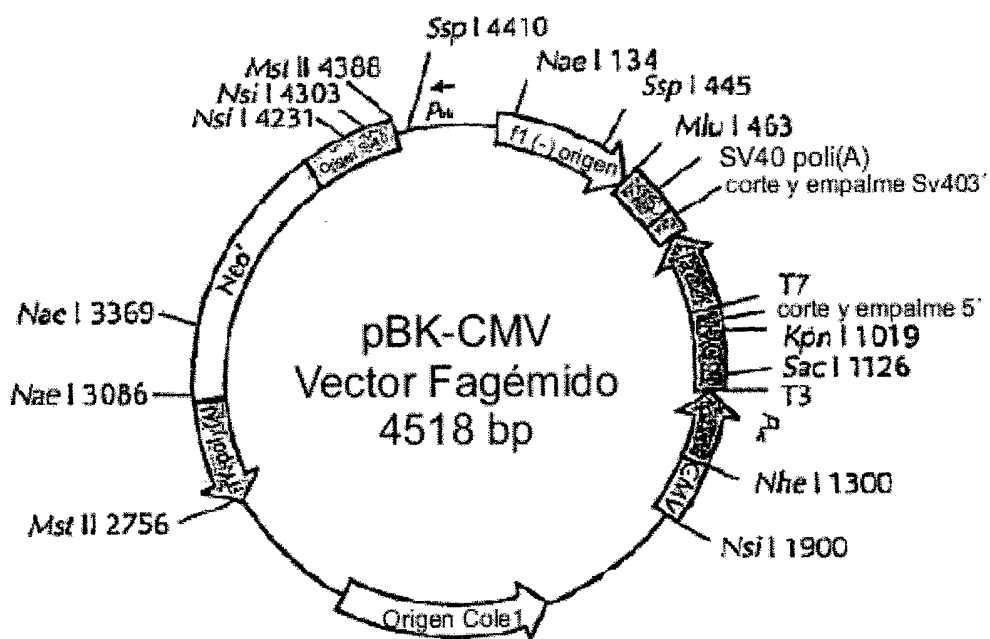
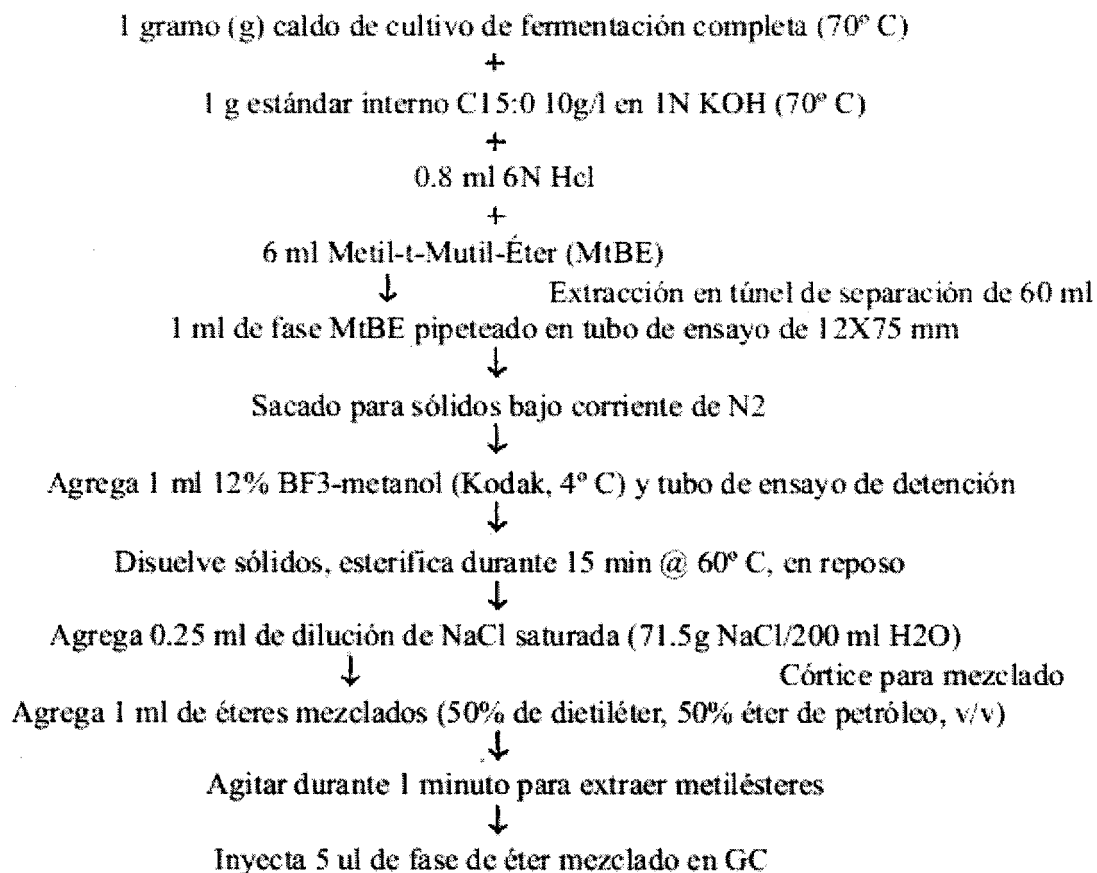


FIG. 17B



Parámetros GC

Columna: columna de capilaridad de HP-INNOWAX, 30m X 0.32 mm, espesor de película 0.5 μ m
 Proporción de división de :1:100
 Presión de Cabeza de Columna: 13.5 psig
 Temperatura de inyector: 240° C
 Temp de Detector FID.: 250° C
 Prog. De Temp.: 90° C durante 0 min. a 190° C @7° C/min. durante min a 235° C @ 12° C/min durante 30 min.

FIG. 18

Origin ColE1: bases 3113-3786



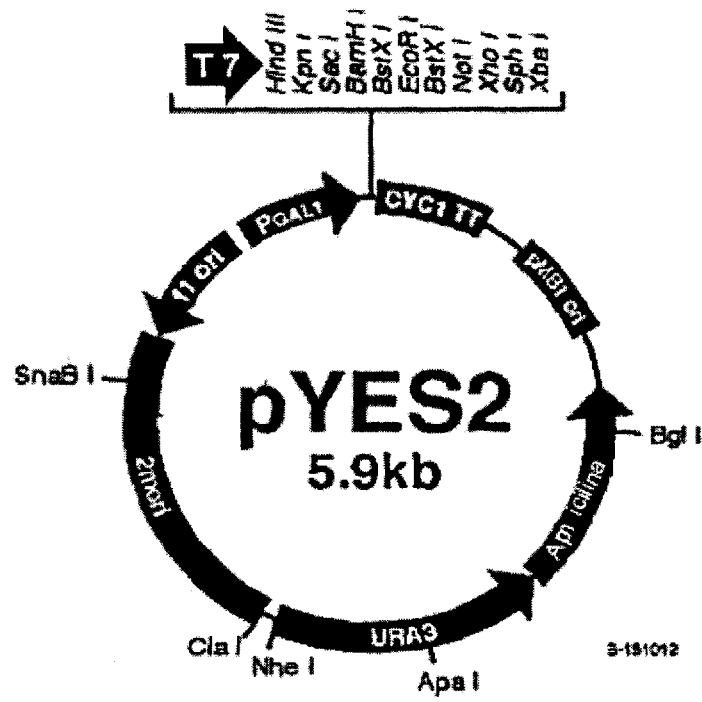


FIG. 20

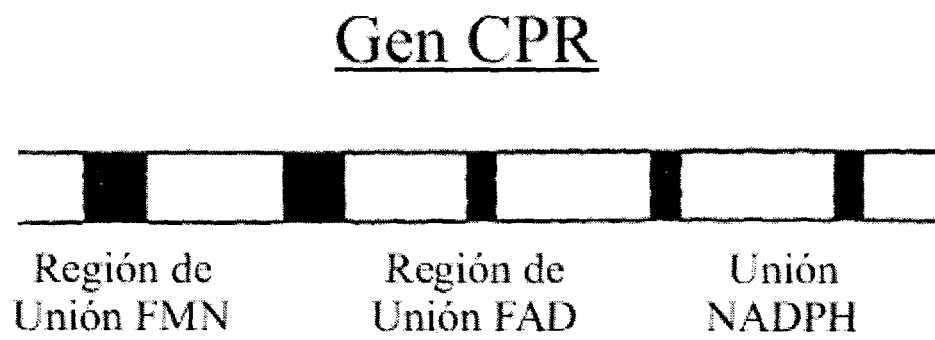


FIG. 21

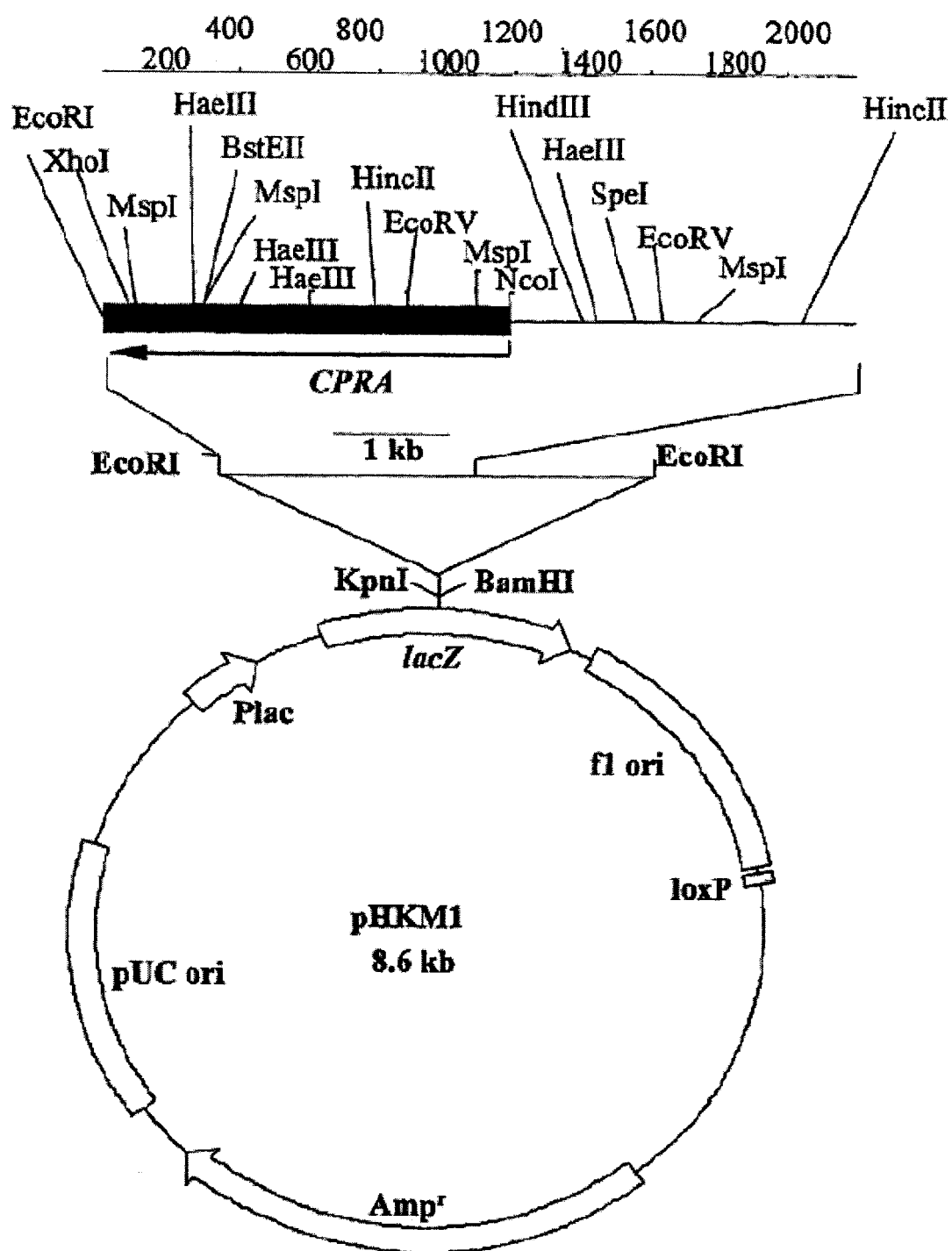


FIG. 22

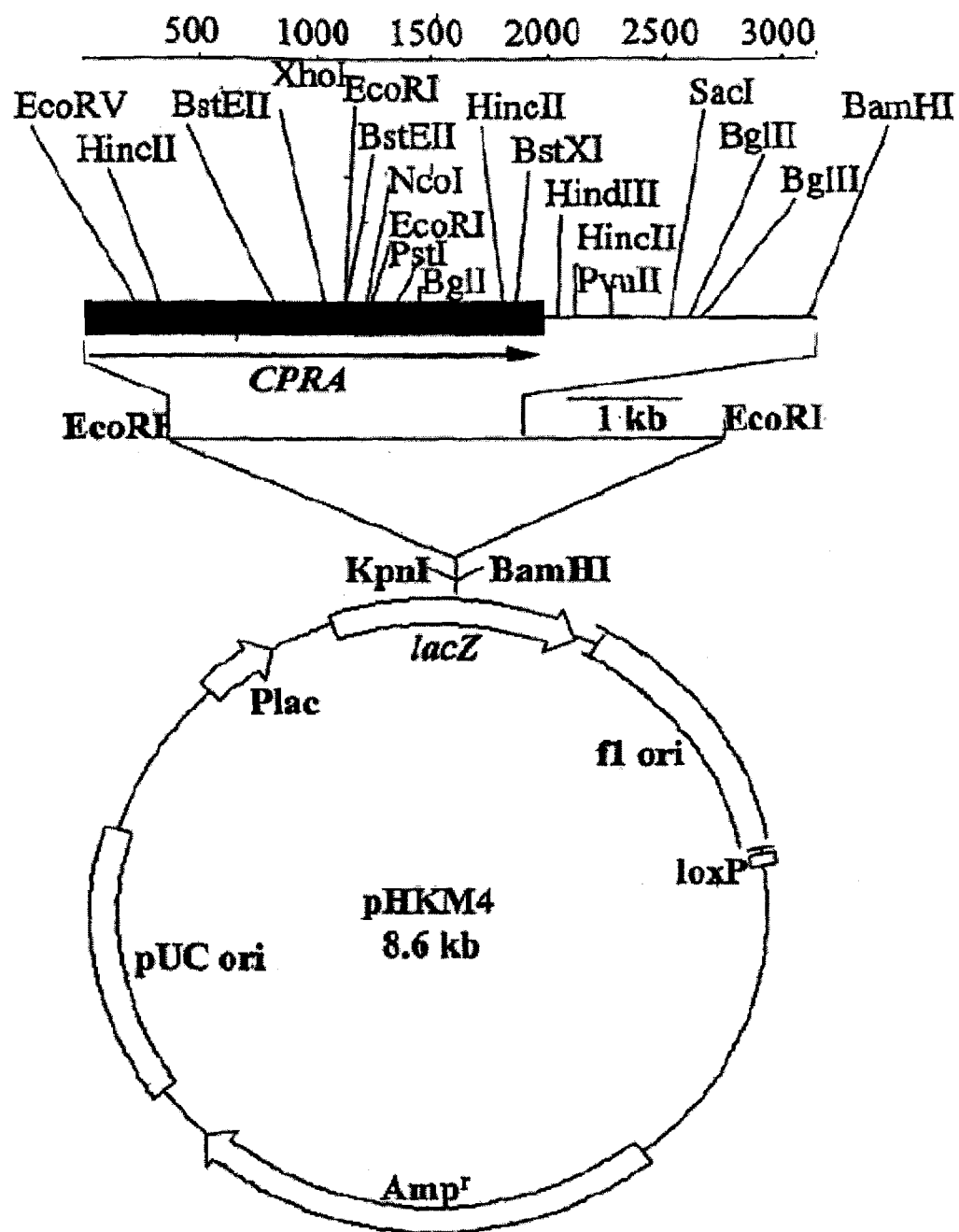


FIG. 23

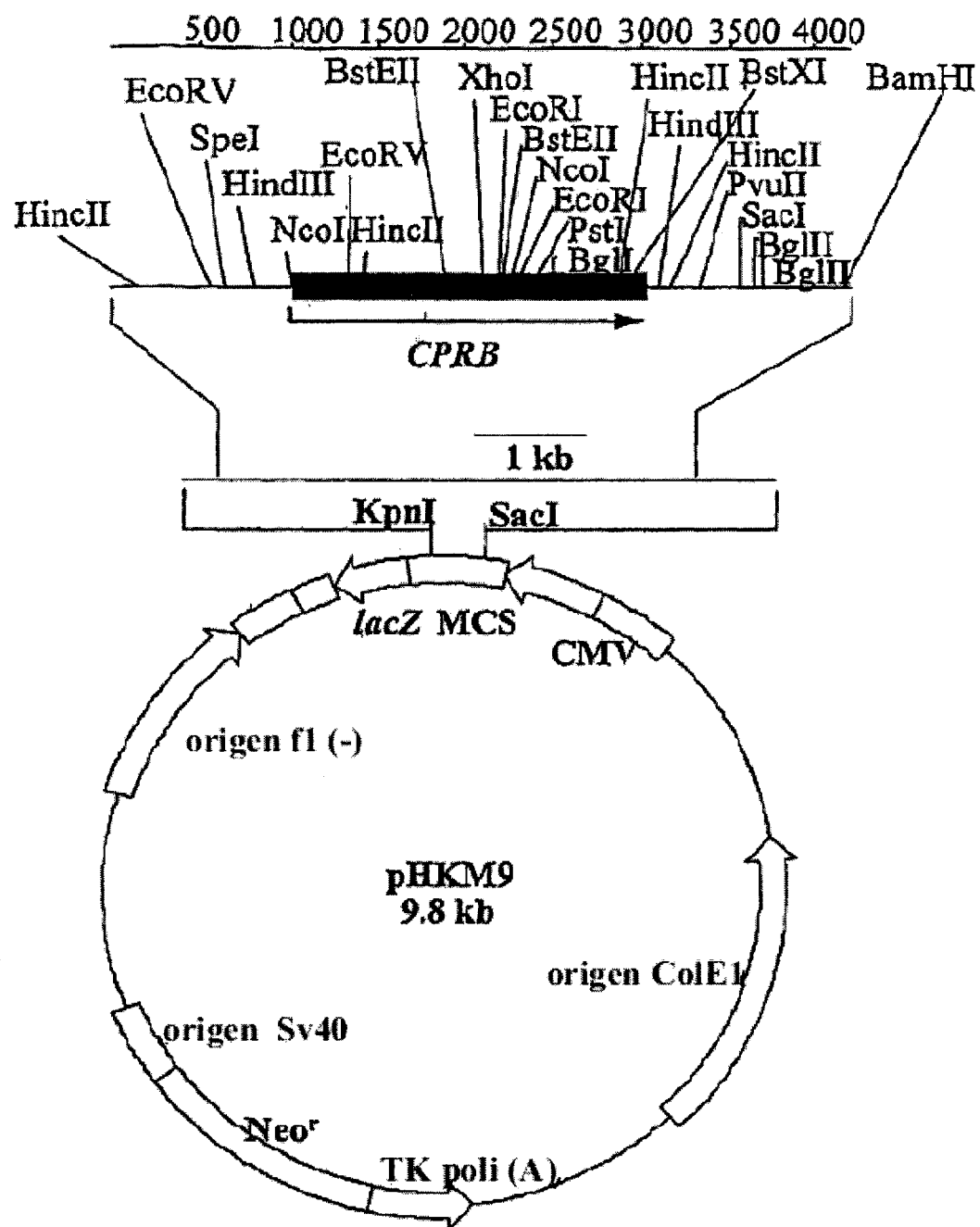


FIG. 24

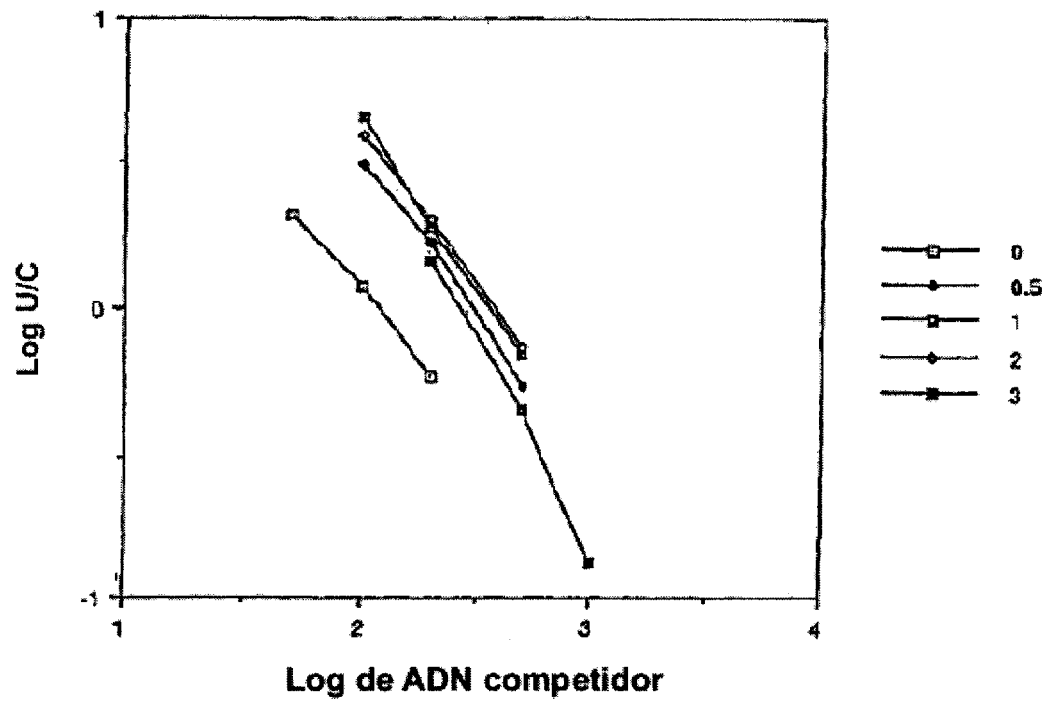


FIG. 25

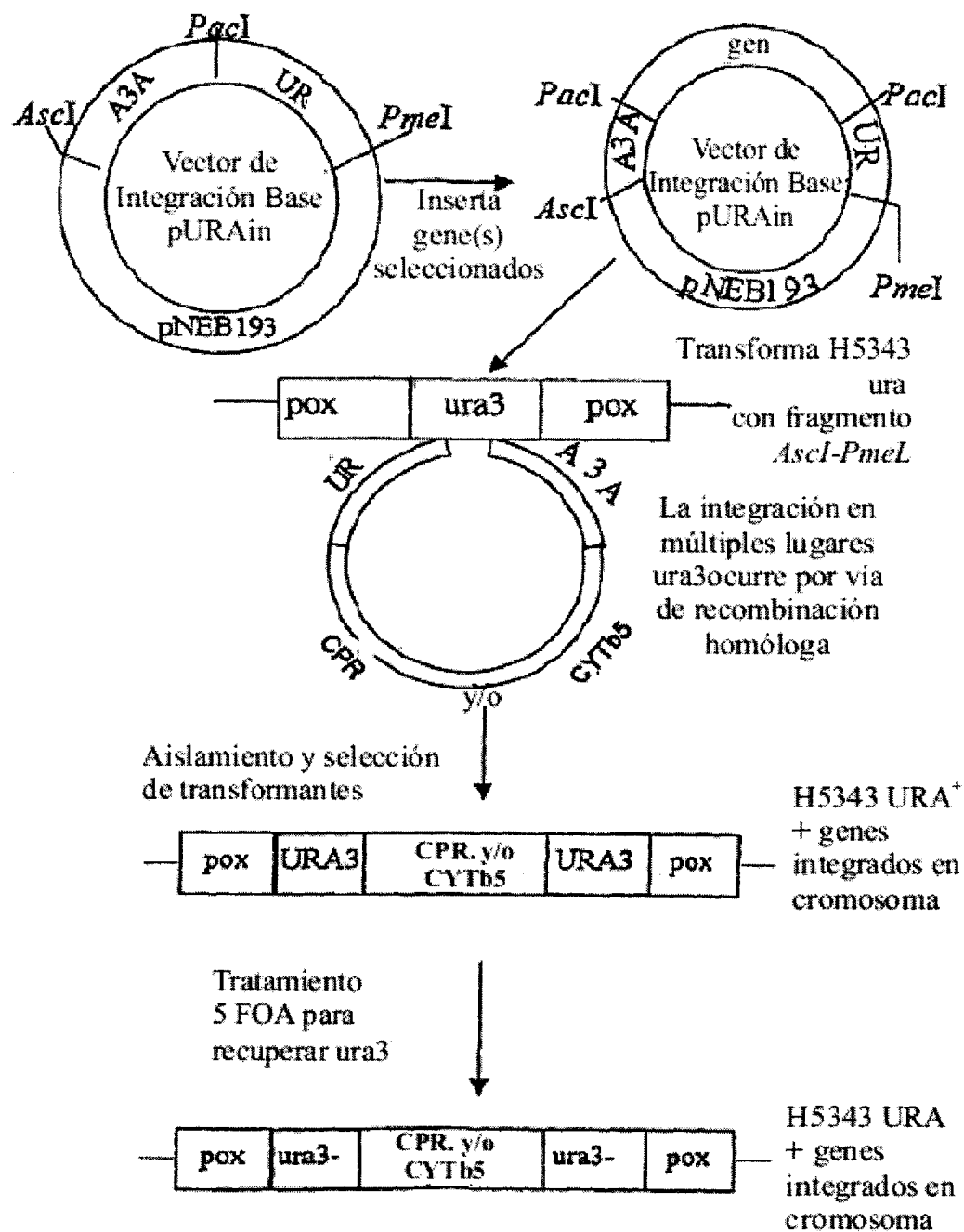


FIG. 26

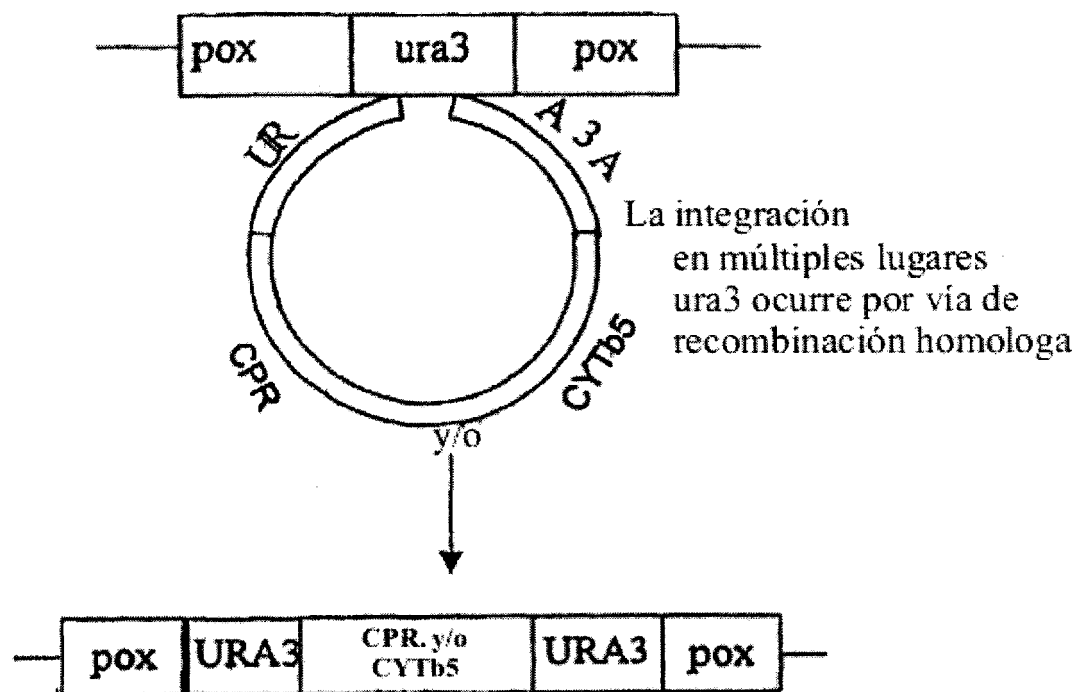


FIG. 27

Rango de Secuencia 1 a 1712

```

10      20      30      40      50      60      70      80      90     100
CGTACCGAGC TCACGACTTT TGGGATTTTC GAGCTTGGAT TGTTCCTT GTTATTTGAA TTGACGAAAC CAGAGCTTTT CAGACAGAT AAGATTGGCT

110     120     130     140     150     160     170     180     190     200
TTATCAAAAC CCACTTTGAA ATATTCAGAT TCGTTTCGAA SATATCTTCA AGAAGATTGA CGATTTGAAA TTGGAAGAG TGGAGAAGAT CTGCTTTGAA

210     220     230     240     250     260     270     280     290     300
TTGTGGAGA ATTTCAGAA TCTCAAGAT TACTCTAAGC ACGGTACAA CGAAGATTGT ATTGAATTGA TCAGCAACAT GATCTTGGTG TTACAGAAAC

310     320     330     340     350     360     370     380     390     400
TCAAGTTCTT GGACCAAGCT GAGAATGCCA CAGATATACA AGGCGTCAAT TGAATAATG GATGAGATT TATCCACAT TGAAGAGAGA GTTTATGCCA

410     420     430     440     450     460     470     480     490     500
AGTGCTCAAC CAGAAGCTAA ACAGGAAGAA GCAAGCAAG AGGTGAACA AGAAGAGAA GSTAATAAG TATTTGTAT TATATAACA ACAAGTAGG

510     520     530     540     550     560     570     580     590     600
GAATTCAGAT TTATACATA AATTCGATA CTAGTCAGT GAGATATCT ATCATTTCC CAATCCCAA GAAAAAAA AAGTAAAAA AAAATCAAA

610     620     630     640     650     660     670     680     690     700
CCCAAGATC AACCTCCCA TCATCATCT CATCAAGCC CCAGCTCAA TCGAATGGT TAGCAGAAA ACATACAG AGAGGATC AGCACCCC

710     720     730     740     750     760     770     780     790     800
TCCAGCTTG CCCAGCTTT ATTCTGTTA ATGAGTCCA AAAAGAGAA CCTTCCGC TCGATGAGC TGACCACAC CCGCAGTTC GTTCTGCTA

810     820     830     840     850     860     870     880     890     900
TCGCAAGCT CCGTCCGAC ATCTGTCTC TGAAGAGCA CATGATATC ATCTAGAT TACCTACGA GGGCAGATT GAGCGTTGC TTGTCTTTC

910     920     930     940     950     960     970     980     990     1000
AGAGCCGAC GGTCTTGA TATTCAGGA CAGGAAGTT CTGATATC GAAACCTG GATCTTCAG TACCTCTGG GGTATACCG GATCCGGCG

1010    1020    1030    1040    1050    1060    1070    1080    1090    1100
TGGACTGACA TCGGAGCC CACCGAGTC ACTGGGAGG CCTCTCTGA AGGTTGAAA CCGCTCCCG AGGCGTACA AAGGAGGG CCGCTGTTA

1110    1120    1130    1140    1150    1160    1170    1180    1190    1200
TGTTCGCGA GTTCTCAGT AAAGCTCT TGGCCATGG TGAATATCC CGTGAACA TCGAGATTG GAGAGTGT CCGGAGTTC TGATTGGTT

1210    1220    1230    1240    1250    1260    1270    1280    1290    1300
CATCGCCAC CCGACATGG GCGGTAGAG AGAAGGCTT GATTGATCA TCATACCC TCGTGTGGG TTGGATGATA AAGCGATG GTTGGCCAG

1310    1320    1330    1340    1350    1360    1370    1380    1390    1400
CACTATAGA CTGTTGATG GGTGGTCTG ACTGATCC AGTGATTAT TGTGGGAA GGTTCCTT GAAAGGAG AGCCCTGAG GTGAGGGA

1410    1420    1430    1440    1450    1460    1470    1480    1490    1500
AGAGATACAC GGATGCTGA TCGAAGCAT ACTGAAGAG AACTGCTAC TTAGAATA TATTCTATA AATAGCTTA TATACATACA CTAACTTCT

1510    1520    1530    1540    1550    1560    1570    1580    1590    1600
AGGAGTCAAT TGTAGTCTT GAAGTGTCT GCTAGTTAG TCTCATGAT TCGAAGCC AATAAGCAA TGGATGAGC AGGATGCT GTTACTGCT

1610    1620    1630    1640    1650    1660    1670    1680    1690    1700
TCTTGAAAA CCGAGATAC GCGCCTCAA ACCAGTCA ATTCGCTTT TGTTCATCC GATCAGCTT GTTGAGGTA TCACGTAGC AGTTGTAATA

1710
CACCTTGAAG AA

```

FIG. 28

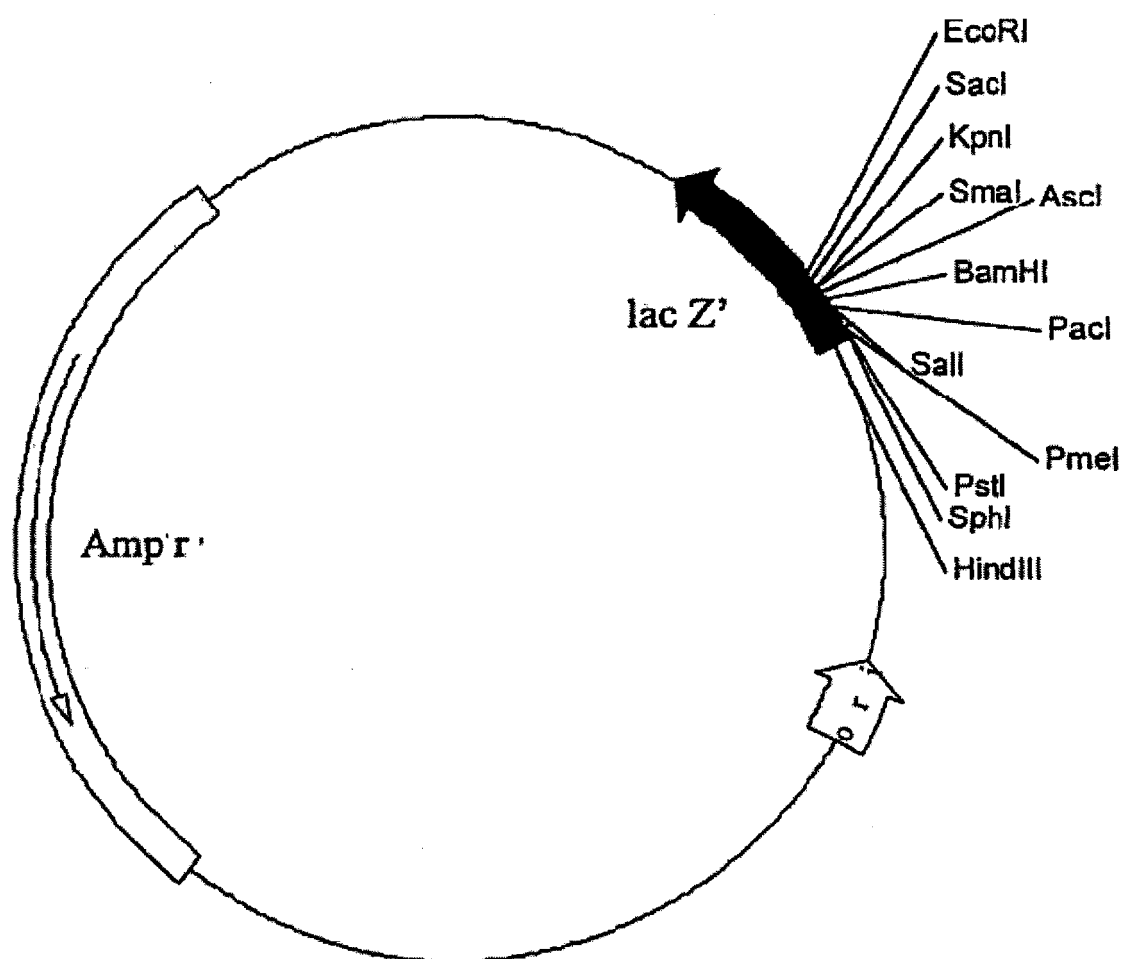


FIG. 29

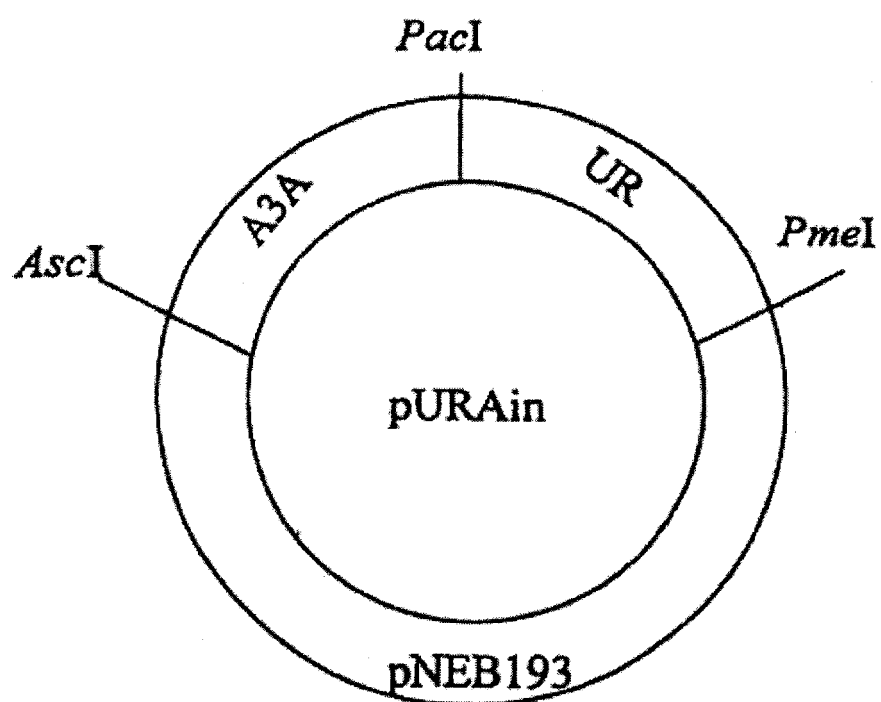


FIG. 30

ES 2 298 276 T3

LISTA DE SECUENCIAS

<110> David, Craft L.
Madduri, Krishna
Loper, John C.

<120> GEN b5 DE CITOCROMO Y PROTEÍNAS DE *CANDIDA TROPICALIS* Y PROCEDIMIENTOS CORRESPONDIENTES

<130> M6368 (1010-35)

<160> 45

<170> PatentIn version 3.1

<210> 1

<211> 2710

<212> ADN

<213> *Candida tropicalis*

<400> 1

```

acataacttca agcagtttgg cgacatagtg aacctcaagt tatcacggaa caagacgacg      60
ggcaagagca agcactacgg gtttatagag ttcacgtcgc ctgaagttgc ccagatcgcg      120
gcggagacga tgaacaacta cttgttggtt ggacacttga tcaaatgtga ggttgtcagc      180
gagccgttca aggacttggt caaggactcg aagaggaagt tcaaggatgat tccctggaag      240
aagatcgcgga aggataagca cgataagcca aagtccgcga aggagtgggc gaagttggtg      300
gagaagttcg aagagtccaa gaagaagaag caggaggagt tgaagagtaa aggtattgat      360
tttgatttgg ctgctatata aaggagataa gagaggagga tgacaagcgc aaacgagcat      420
tctgttgatg tgtaagcag gtatagataa tagggataa cgtaaaataa gagatctcca      480
acttccaaact tccaacttcc gacctcctc ttttggggga gagggttggt tatgtagtgg      540
tgaggagagag gaggatattt tgttttgctt aattgggata aattatccca gtcagttgaa      600
agagcgaggg gtaagccatt tctttttcta actgcaata gcatacagat gcgatatgta      660
acgaagagag aatcaagag caggtgacta catacataga tagtgacatt ataataacat      720
ggcgcatcat tggttctatg tagctggcag gggtattatc aagcttgaat agtttaataa      780
aaatcgtacc atgaatgtat gcatagaagc aataaggaag cctgtgcctg tgagtagtag      840
cagtagcggg gggagacgct agtttagggg taaaatgtca gcacatgaac agcagttgaa      900
gtgggtgcc aatcaagtaag aacatcttgt gaaaaatcaa aagcaatggt atatgtgttc      960
ctgcatacag tgctggagtc aacgagccaa aaaaaaaaaa gaaagaaaga gageaaaact     1020
tatcgtataa aaaccacaca aaaatttccc aatcccaatt ccttcattct tcttctttta     1080
ctgatttaac ccacagatac atacaattat gaccgacaca gacaccacga ccaccatcta     1140
caccacgaa gaggttgccc agcacaccac ccacgacgac ttgtgggtta ttctcaatgg     1200
taaggtctac aacatctcca actatataga cgagcaccga ggtggtgaag aagtcattct     1260

```

ES 2 298 276 T3

5 tgattgcgcc ggcacagacg ccactgaagc ctttgacgac attggccact ccgacgaggc 1320
 ccacgagatc ttggaaaagt tgtacattgg taacttgaag ggcgctaaga ttgttgaggc 1380
 10 caagcacgcg cagtcgttca gcacggnga agactcgggt atcaacttcc cattgattgc 1440
 tgttggtgtg tttttggctg ctttcgggtg ctactactac aagaccaact ttgcctaacc 1500
 ataacaagca gtacagttga aggacagggt agaggagatg agaaaaaacg ggaacccaac 1560
 aaagattatt ttcacacatc acatggaggg gctgatccca ctttttgacg tcaatatcca 1620
 cagcacgaag aaagaaagaa agaaagaaag totatggaag aggaatgga tcacattaga 1680
 15 ccttttcttt atgtaacata tatatatata taaactaata cagatttaca gatacaccac 1740
 atcacgcgag ggcttatcat ctgatggtgc ccaaaaaaaaa aaatccactg tggatgagcc 1800
 tagttaggag atatcggagt agtcattctt tttgatctt aggtcttctt ctcttgatt 1860
 20 ctacgttggc acttggtgct acacgatgag atcaccagggt gtcattctgg agtttggtgg 1920
 aaagtgtgtt gattttttta gtaagcaaga atttgttgag ttctattgga tgttctggtg 1980
 cggcgaactc catcccccca ccccttgctt tgtcttctt tgtcttctt ttttgggtcg 2040
 25 gttggcgga gtaagacgca cgcacaggag gagcacgacg gataaatatc cacttttttc 2100
 acacgcgtcg attgacggct tgtgtgaatt gtggggaata cggataaggg ggtataccac 2160
 acacacacat atctaacata tcagaccact ttctataaca gatctcatga tcccttgag 2220
 agttgatgca agtctatgct cctgtgatat tgcctctctt ccccpaagga agggcggggc 2280
 atgttatcag ggacctgat gaaccttga tggcgggtgt agtagatgca agagaggttg 2340
 35 tgcttttgaa gtagctgaag gtgtagggac atccggtact atagttctct tgaaggatca 2400
 tgccagctcc ctttctgtgg ctctctggaa gctctgcato ttctcttctt tgaaacagcg 2460
 tggagttacg aaaggtaacc tgtggtgagt tcaaacaaga catggctcta caagctgtcg 2520
 40 aggataaaag taattaaaca acatgtatat atattaataa acggatccgt ggtgctagat 2580
 tgtggtagat gtttagtato gtttatcacc totagtgaag actagccttt gattcctta 2640
 gtcatcagta cttgatgta cattcaacca aatgaaggtc ggtccaagat ccaaagaatt 2700
 45 caaaaagctt 2710

50 <210> 2
 <211> 129
 <212> PRT
 <213> *Candida Tropicalis*

55

60

65

ES 2 298 276 T3

<400> 2

5 Met Thr Asp Thr Asp Thr Thr Thr Thr Ile Tyr Thr His Glu Glu Val
1 5 10 15

Ala Gln His Thr Thr His Asp Asp Leu Trp Val Ile Leu Asn Gly Lys
20 25 30

10 Val Tyr Asn Ile Ser Asn Tyr Ile Asp Glu His Pro Gly Gly Glu Glu
35 40 45

15 Val Ile Leu Asp Cys Ala Gly Thr Asp Ala Thr Glu Ala Phe Asp Asp
50 55 60

20 Ile Gly His Ser Asp Glu Ala His Glu Ile Leu Glu Lys Leu Tyr Ile
65 70 75 80

Gly Asn Leu Lys Gly Ala Lys Ile Val Glu Ala Lys His Ala Gln Ser
85 90 95

25 Phe Ser Thr Glp Glu Asp Ser Gly Ile Asn Phe Pro Leu Ile Ala Val
100 105 110

30 Gly Val Phe Leu Ala Ala Phe Gly Val Tyr Tyr Tyr Lys Thr Asn Phe
115 120 125

35 Ala

<210> 3

<211> 4206

40 <212> ADN

<213> *Candida tropicalis*

<400> 3

45 catcaagatc atctatgggg atgattacga cagcaacatt gcagaaagag cgttggtcac 60

aatcgaaaga goctatggcg ttgcogtcgt tgaggcaaat gacagcacca acaataacga 120

50 tgggccaggt gaagagcctt cagaacagtc cattgttgac gcttaaggca cggataatta 180

cgtggggcaa aggaacgcgg aattagttat ggggggatca aaagcggaag atttgtgttg 240

55 cttgtgggtt ttttccttta ttttccatat gattttcttg cgcaagtaac atgtgccaat 300

ttagtttgtg attagcgtgc ccacaattg gcacgttga cgggcgtgtt ttgtcatacc 360

ccaagtctta actagctcca cagtctcgac ggtgtctega cgtgtcttc ttcacccct 420

60 cccatgaatc attcaaagt gtgggggat ctcaccaag ggcacggag ttaetgctta 480

tgtttctccc actttggttg tgattgggtt agtctagtga gttggagatt ttcttttttt 540

65

ES 2 298 276 T3

	gatggcagac cagtcactgg tgttgcacc aacttggtga agaacggtga aattgtgcaa	2460
	aacaagactg gcgaaaagcc acttgtccac tacgatttga gcggcccaag aggcaagttc	2520
5	aacaagttca agttgceagt gcattgtgaga agatccaaact ttaagttgoc aaagaactcc	2580
	accaccccag ttatcttgat tgggccaggt actggtgttg cccattgag aggtttgtc	2640
10	agagaaagag ttcaacaagt caagaatggt gtcaatgttg gcaagacttt gttgttttat	2700
	ggttgcagaa actccaacga ggactttttg tacaagcaag aatgggccga gtacgcttct	2760
	gttttgggtg aaaactttga gatgttcaat gccttctcca gacaagacc atccaagaag	2820
15	gtttacgtcc dggataagat tttagaaaac agccaacttg tgcacgagtt gttgactgaa	2880
	ggtgccatta tctacgtctg tgggtatgcc agtagaatgg ctagagacgt gcagaccaca	2940
	atttccaaga ttgttgctaa aagcagagaa attagtgaag acaaggctgc tgaattggtc	3000
20	aagtccctga aggtcgaasa tagataccaa gaagatgttt ggtagactca aacgaatctc	3060
	tctttctccc aacgcattta tgaatcttta ttctcattga agctttacat atgttctaca	3120
	ctttattttt tttttttttt ttattattat attacgaaac atagggtcaac tataataact	3180
25	tgattaaatg ttatagaaac aataactatb atctaactgt ctacttcttt ggcattgaca	3240
	tcaacattac cgttccatt accgttgccg ttggcaatgc cgggatattt agtacagtat	3300
30	ctccaatccg gatttgagct attgtagatc agctgcaagt cattctccac cttcaaccag	3360
	tacttatact tcatctttga cttcaagtcc aagtcataaa tattacaagt tagcaagaac	3420
	ttctggccat ccaogataata gacgttatcc acgttatiat gcgaogtatg gatgtggtta	3480
35	tcottattga acttctcaaa cttcaaaaac aaccccaagt cccgcaactg cattatcaac	3540
	gacaagttct ggctcagtc gtccgagctc gtcaagttct caattagatc gttcttgtta	3600
40	ttgatcttct ggtactttct caattgctgg aacacattgt cctcgttgtt caaatagatc	3660
	ttgaacaact ttttcaacgg gatcaacttc tcaatctggg ccaagatctc cgcggggtc	3720
	ttcagaaaca agtccgtcaa cccctggctg atggtctccg ggtacaacaa gtccaagggg	3780
45	cagaagtgtc taggcacgtg ttccaactgg ttcaacgaac atgttcgaca gtagttcgag	3840
	ttatagttat cgtacaacca ttttggtttg atttcgaaa tgacggagct gatgocatca	3900
	ttctcctggt tcctctcata gtacaactgg cacttcttcg agaggtctca ttctcgtag	3960
50	ttcccggtcca agatattcgg caacaagagc ccgtaccgct caccggagcat caagtcgtgg	4020
	ccctggttgt tcaacttgtt gatgaagtc gaggtcaaga caatcaactg gatgtcgatg	4080
55	atctggtgcg ggaacaagtt cttgcatttt agctcgatga agtcglacaa ctccacgctc	4140
	gagatatact cctgttcttc cttcaagagc cggatccgca agagcttgtg cttcaagtag	4200
	tcgttg	4206

60

<210> 4

<211> 679

65 <212> PRT

<213> *Candida tropicalis*

ES 2 298 276 T3

<400> 4

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

ES 2 298 276 T3

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

ES 2 298 276 T3

Ala Asp Gly Arg Pro Val Thr Gly Val Val Thr Asn Leu Leu Lys Asn
 465 470 475 480
 5 Val Glu Ile Val Gln Asn Lys Thr Gly Glu Lys Pro Leu Val His Tyr
 485 490 495
 10 Asp Leu Ser Gly Pro Arg Gly Lys Phe Asn Lys Phe Lys Leu Pro Val
 500 505 510
 15 His Val Arg Arg Ser Asn Phe Lys Leu Pro Lys Asn Ser Thr Thr Pro
 515 520 525
 20 Val Ile Leu Ile Gly Pro Gly Thr Gly Val Ala Pro Leu Arg Gly Phe
 530 535 540
 25 Val Arg Glu Arg Val Gln Gln Val Lys Asn Gly Val Asn Val Gly Lys
 545 550 555 560
 30 Thr Leu Leu Phe Tyr Gly Cys Arg Asn Ser Asn Glu Asp Phe Leu Tyr
 565 570 575
 35 Lys Gln Glu Trp Ala Glu Tyr Ala Ser Val Leu Gly Glu Asn Phe Glu
 580 585 590
 40 Met Phe Asn Ala Phe Arg Gln Asp Pro Ser Lys Lys Val Tyr Val
 595 600 605
 45 Gln Asp Lys Ile Leu Glu Asn Ser Gln Leu Val His Glu Leu Leu Thr
 610 615 620
 50 Glu Gly Ala Ile Ile Tyr Val Cys Gly Asp Ala Ser Arg Met Ala Arg
 625 630 635 640
 55 Asp Val Gln Thr Thr Ile Ser Lys Ile Val Ala Lys Ser Arg Glu Ile
 645 650 655
 60 Ser Glu Asp Lys Ala Ala Glu Leu Val Lys Ser Trp Lys Val Gln Asn
 660 665 670
 65 Arg Tyr Gln Glu Asp Val Trp
 675

<210> 5

<211> 4145

<212> ADN

<213> *Candida tropicalis*

ES 2 298 276 T3

<400> 5

	tatatgatat atgatatatc ttccctgtgta attattattc gtatttcgtta atacttacta	60
5	cattttttttt tctttattta tgaagaaaag gagagttcgt aagttgagtt gagtagaata	120
	ggctgttggtg catacgggga gcagaggaga gtatccgacg aggagggaact ggggtgaaatt	180
10	tcatttatgc tgttgcgcc tgtactgtac tgtaaattct agatttccta gaggttgctc	240
	tagcaaataa agtgttticaa gatacaattt tacaggcaag ggtaaaggat caactgatta	300
15	gcggaagatt ggtgttgctt gtgggggtctt tttatttttc atatgatttc tttgcgcgag	360
	taacatgtgc caactatggt tatgattagc gtacctccac aattggcacc ttggacgggc	420
	gtgttttgtc ttaacccaag ccttatttag ttccacagtc tcgacgggtg ctgcgcgag	480
20	tcttctccca cccctcgag gaatcattcg aagttgttg gggatctct cgcagttta	540
	tggtcatgtc tttcccaatt tgggtgtgat tggggtagcg tagtgagttg gtgattttct	600
25	tttttcgcag gtgtctccga tatcgaagtt tgatgaatat aggagccaga tcagcatggt	660
	atattgcctt ttagataga gatgttgac acaactagc tgaattacac accaccgcta	720
	aacgatgcgc acaggggtgc accgccaaact gacgttgggt ggagttgttg ttggcagggc	780
30	catattgcta aacgaagaga agtagcaca aaccaaggt taagaacaat taaaaaatt	840
	catacgaca ttccacagcc atttacataa tcaacagcga caaatgagac agaaaaaact	900
35	ttcaacattt caaagttccc tttttcttat tacttctttt tttotttctt tcttttcatt	960
	tcttttctt ctgcttttat tactttacca gtottttgct tgtttttgca attcctcatc	1020
	ctctctctca ccatggcttt agacaagtta gatttgtatg tcatcataac attggttggtc	1080
40	gctgtggccg cctattttgc taagaaccag ttcttgatc agccccagga caccgggttc	1140
	ctcaacacgg acagcgggaag caactccaga gacgtcttgc tgacattgaa gaagaataat	1200
45	aaaaacacgt tgttgttgtt tgggtcccag accggtacgg cagaagatta cgccaacaaa	1260
	ttgtcaagag aattgcactc cagatttggc ttgaaaacca tggttgcaga tttcgtgat	1320
	tacgattggg ataacttogg agatatcacc gaagatatct tgggtgtttt catcgttgcc	1380
50	acctacggtg aggggtgaacc taccgacaat gccgacgag tccacacctg gttgactgaa	1440
	gaagctgaca ctttgagtac tttgagatat accgtgttcg ggttgggtaa ctccacctac	1500
55	gagttcttca atgctattgg tagaaagttt gacagattgt tgagtgagaa aggtggtgac	1560
	agatttgctg aatatgctga aggtgacgac ggcactggca ccttggacga agatttcag	1620
	gcctggaagg ataattgtct tgacgccttg aagaatgact tgaactttga agaaaaggaa	1680
60	ttgaagtacg aaccaaactg gaaattgact gagagagatg acttgtctgc tgccgactcc	1740
	caagtttctt tgggtgagcc aaacaagaag tacatcaact ccgagggcat cgacttgacc	1800
65	aaggggccat tcgaccacac ccaccatac ttggccagga tcaccgagac cagagagttg	1860

ES 2 298 276 T3

ttcagctcca aggaaagaca ctgtattcac gttgaatttg acatttctga atcgaacttg 1920
 aaatacacca ccggtgacca tctagccatc tggccatcca actccgacga aaacatcaag 1980
 5 caatttgcca agtggttcgg attggaagat aaactcgaca ctgttattga attgaaggca 2040
 ttggactcca cttacaccat tccattccca actccaatta cttacgggtgc tgtcattaga 2100
 10 caccatttag aaatctccgg tccagtctcg agacaattct ttttgcgat tgcgtgggtt 2160
 gctcctgatg aagaaacaaa gaagactttc accagacttg gtggtgacaa acaagaattc 2220
 gccaccaagg ttaccgcgag aaagttcaac attgccgatg ccttggtata ttctccaac 2280
 15 aacactccat ggtccgatgt tccttttgag ttcttattg aaaacatcca acacttgact 2340
 ccaogttact actccatttc ttcttcgtcg ttgagtga aaacaactcat caatgttact 2400
 20 gcagtcgttg aggcggaaga agaagccgat ggcagaccag tcaactggtg tgttaccaac 2460
 ttgttgaga acattgaaat tgcgcaaac aagactggcg aaaagccact tgttactac 2520
 gatttgagcg gcccaagagg caagttcaac aagttcaagt tgccagtga cgtgagaaga 2580
 25 tccaacttta agttgcaaaa gaactccacc accccagtta tcttgattgg tccaggtact 2640
 ggtgttgccc cattgagagg ttctggtaga gaaagagttc aacaagtcaa gaatggtgtc 2700
 30 aatgttgcca agactttggt gttttatggt tgcagaaact ccaacgagga cttttgtac 2760
 aagcaagaat gggcgagta cgttctgtt ttgggtgaaa actttgagat gttcaatgcc 2820
 ttctctagac aagaccatc caagaaggtt tacgtccagg ataagatttt agaaaacagc 2880
 35 caacttgtgc cgaatttgtt gaccgaaggt gccattatct acgtctgttg tgacgcaggt 2940
 agaatggcca gagacgtcca gaccacgatc tccaagattg ttgccaaaag cagagaaatc 3000
 40 egtgaagaca aggcgctga atttggtcaag tcttggaag tccaaatag ataccaagaa 3060
 gatgtttggt agactcaaac gaatctctct ttctcccaac gcatttatga atattctcat 3120
 tgaagtttta catatgttct atatttcatt tttttttat tatattacga aacataggctc 3180
 45 aactatatat acttgattaa atgttataga aacaataatt attatctact cgtctacttc 3240
 tttggcattg gcattggcat tggcattggc attgccgttg ccgttggtaa tgccgggata 3300
 50 tttagtacag tatctccaat ccgatttga gctattgtaa atcagctgca agtcattctc 3360
 caccotcaac cagtacttat acttcatctt tgacttcaag tccaagtcac aaatattaca 3420
 agttagcaag aactctggc catccacaat atagacgtta ttcacgttat tatgcgacgt 3480
 55 atggatatgg ttaiccttat tgaacttctc aaacttcaaa aacaaccca cgtcccgcaa 3540
 cgtcattatc aacgacaagt tctgactcac gtcgtcggag ctggtcaagt tctcaattag 3600
 60 atcgttcttg ttattgatct tctggtactt tctcaactgc tggaaacacat tgcctcgtt 3660
 gttcaaatag atcttgaaca acttcttcaa gggaatcaac ttttcgatct gggccaagat 3720

65

ES 2 298 276 T3

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

ES 2 298 276 T3

Phe Phe Leu Ser Ile Ala Gly Phe Ala Pro Asp Glu Glu Thr Lys Lys
 370 375 380
 5
 Thr Phe Thr Arg Leu Gly Gly Asp Lys Gln Glu Phe Ala Thr Lys Val
 385 390 395 400
 10
 Thr Arg Arg Lys Phe Asn Ile Ala Asp Ala Leu Leu Tyr Ser Ser Asn
 405 410 415
 15
 Asn Thr Pro Trp Ser Asp Val Pro Phe Glu Phe Leu Ile Glu Asn Ile
 420 425 430
 20
 Gln His Leu Thr Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Ser Leu Ser
 435 440 445
 25
 Glu Lys Gln Leu Ile Asn Val Thr Ala Val Val Glu Ala Glu Glu Glu
 450 455 460
 30
 Ala Asp Gly Arg Pro Val Thr Gly Val Val Thr Asn Leu Leu Lys Asn
 465 470 475 480
 35
 Ile Glu Ile Ala Gln Asn Lys Thr Gly Glu Lys Pro Leu Val His Tyr
 485 490 495
 40
 Asp Leu Ser Gly Pro Arg Gly Lys Phe Asn Lys Phe Lys Leu Pro Val
 500 505 510
 45
 His Val Arg Arg Ser Asn Phe Lys Leu Pro Lys Asn Ser Thr Thr Pro
 515 520 525
 50
 Val Ile Leu Ile Gly Pro Gly Thr Gly Val Ala Pro Leu Arg Gly Phe
 530 535 540
 55
 Val Arg Glu Arg Val Gln Gln Val Lys Asn Gly Val Asn Val Gly Lys
 545 550 555 560
 60
 Thr Leu Leu Phe Tyr Gly Cys Arg Asn Ser Asn Glu Asp Phe Leu Tyr
 565 570 575
 65
 Lys Gln Glu Trp Ala Glu Tyr Ala Ser Val Leu Gly Glu Asn Phe Glu
 580 585 590
 Met Phe Asn Ala Phe Ser Arg Gln Asp Pro Ser Lys Lys Val Tyr Val
 595 600 605
 Gln Asp Lys Ile Leu Glu Asn Ser Gln Leu Val His Glu Leu Leu Thr
 610 615 620

ES 2 298 276 T3

Glu Gly Ala Ile Ile Tyr Val Cys Gly Asp Ala Ser Arg Met Ala Arg
625 630 635 640

5 Asp Val Gln Thr Thr Ile Ser Lys Ile Val Ala Lys Ser Arg Glu Ile
645 650 655

10 Ser Glu Asp Lys Ala Ala Glu Leu Val Lys Ser Trp Lys Val Gln Asn
660 665 670

Arg Tyr Gln Glu Asp Val Trp
675

15

<210> 7

<211> 4115

20 <212> ADN

<213> *Candida tropicalis*

<400> 7

25

catatgcgct aatcttcttt ttctttttat cacaggagaa actatccac cccacttcg 60

aaacacaatg acaactcctg cgttaacttg aaattcttgt ctgactaatt gaaaactcgg 120

30

gacgagtcag aacctcagtc aaacggacag acagacaaac acttggtgcg atgttcatac 180

ctacagacat gtcaacgggt gtttagacgac ggtttcttgc aaagacaggt gttggcatct 240

cgtacgatgg caactgcagg aggtgtogac ttctcttta ggcaatagaa aaagactaag 300

35

agaacagcgt ttttacaggt tgcattgggt aatgtagtat ttttttagtc ccagcattct 360

gtgggttgct ctgggtttct agaataggaa atcacaggag aatgcaaat cagatggaag 420

40

aacaaagaga taaaaaaca aaaaaaactg agttttgcac caatagaatg tttgatgata 480

tcatecactc gctaaacgaa tcatgtgggt gatcttctct ttagtttttg tctatcataa 540

aacacatgaa agtgaatcc aaatacacta cactccgggt attgtcttc gttttacaga 600

45

tgtctcattg tcttactttt gaggtcatag gagttgctg tgagagatca cagagattat 660

cacactcaca tttatcgtag ttctctatct catgtgtgt gtctctggtt ggttcctgag 720

tttggattgt tgtacattaa aggaatcgct gcaaaagcaa gctaactaaa ttttctttgt 780

50

cacaggtaca ctaacctgta aaacttcaact gccacgcag tctttcctga ttgggcaagt 840

gcacaaacta caacctgcaa aacagcactc cgtttgtcac aggttgtctc ctctcaacca 900

55

acaaaaaat aagattaaac tttctttgt catgcatcaa tcggagttat ctctgaaaga 960

gttgcccttg tgaatgtgt gccaaactca aactgcaaaa ctaaccacag aatgatttcc 1020

ctcacaatta tataaactca cccacatttc cacagaccgt aatttcattg ctcaatttct 1080

60

cttttgctct tcttttactt agtcagggtt gataacttcc tttttatta cccatctta 1140

65

ES 2 298 276 T3

	tttattttatt tattcattta taccaacca ccaaccatgg ccacaccaaga aatcatogat	1200
	tctgtacttc cgtacttgac caaatggtac actgtgatta ctgcagcagt attagtcttc	1260
5	cttatctcpa caaacatcaa gaactacgtc aaggcaaaga aattgaaatg tgtcgatcca	1320
	ccatacttga aggatgcogg tctcactggg attotgtctt tgatcgccgc catcaaggcc	1380
10	aagaacgacg gtagattggc taactttgcc gatgaagttt tcgacgagta cccaaaccac	1440
	accttctact tgtctgttgc cggtgctttg aagattgtca tgactgttga ccagaaaaac	1500
	atcaaggctg tcttggccac ccaattcact gacttctcct tgggtaccag acacgcccac	1560
15	tttgctcctt tgttgggtga cggatctctc accttggacg gagaagggtg gaagcactcc	1620
	agagctatgt tgagaccaca gtttgctaga gaccagattg gacacgttaa agccttggaa	1680
20	ccacacatcc aatcatggc taagcagatc aagttgaacc agggaaagac ttctgatatc	1740
	caagaattgt tctttagatt taocgtcgac aocgctactg agttcttgtt tgggtgaatcc	1800
	gttcaactcc tgiacgatga aaaattgggc atoccaaetc caaacgaaat ccaggaaga	1860
25	gaaaactttg ccgctgcttt caacgtttcc caacactact tggccaccag aagttactcc	1920
	cagacttttt actttttgac caaccctaag gaattcagag actgtaacgc caaggtccac	1980
	cacttggcca agtactttgt caacaaggcc ttgaacttta ctctgaaga actcgaagag	2040
30	aaatccaagt ccggttacgt tttcttgtac gaattgggtt agcaactcag egetccaaag	2100
	gtcttgcaag atcaattgtt gaacattatg gttgcgggaa gagacaccac tgcoggtttg	2160
35	ttgtccittg ctttgtttga atttgctaga caccagaga tgtggtccaa gttgagagaa	2220
	gaaatcgaag ttaacttttg tgttggtgaa gactcccgcg ttgaagaaat taccttcgaa	2280
40	gccttgaaga gatgtgaata cttgaagget atccttaacg aaacotttgc tatgtacca	2340
	tctgttccgt tcaactttag aaccgccacc agagacacca ctttgccaag aggtggtggt	2400
	gctaacggtg ccgacccaat ctacattcct aagggtcca ctgttgctta cgttgtctac	2460
45	aagaoccacc gtttgaaga atactacggt aaggacgcta acgaattcag accagaaaga	2520
	tggtttgaac catctactaa gaagttgggc tgggcttatg ttocattcaa cgggtggtcca	2580
50	agagtctgct tgggtcaaca attcgccttg actgaagett cttatgtgat cactagattg	2640
	gcccagatgt ttgaaactgt ctcatctgat ccaggtctcg aataacctcc accaaagtgt	2700
	attcaactga ccatgagtca caacgatggt gtctttgtca agatgtaaag tagtcgatgc	2760
55	tgggtattcg attacatgtg tataggaaga ttttggtttt ttattcgttc ttttttttaa	2820
	tttttgttaa attagtttag agatttcatt aatacataga tgggtgctat ttccgaaact	2880
	ttacttctat ccctgtatc ccttattatc cctctcagtc acatgattgc tgtaattgtc	2940
60	gtgcaggaca caaactccct aacggactta aaccataaac aagctcagaa ccataagccg	3000

65

ES 2 298 276 T3

acatcaactcc ttctttctctc ttctccaacc aatagcatgg acagacccac cctccctatcc 3060
 gaatcgaaga cctttattga ctccatacc accctggaagc cctcaagcc acacacgtcc 3120
 5 tccagccac ccatcaccac atccctctac tcgacaacgt ccaaagacgg cgagttctgg 3180
 tgtgcccga aatcagccat ccggccaca tacaagcagc cgttgattgc gtgcatactc 3240
 10 ggcgagccca caatgggagc cagcattcg gaccatgaag caaagtacat tcacgagatc 3300
 acgggtggtt cagtgtcga gattgagaag ttcgacgatg gatggaagta cgatctcgtt 3360
 ggcgattacg acttcggtgg gtgtttatct aaacgaagat tctatgagac gcagcatgtg 3420
 15 ttccggttcg aggattgtgc gtacgtcatg agtgtgcctt ttgatggacc caaggaggaa 3480
 ggttacgtgg ttgggacgta cagatccatt gaaaggttga gctggggtaa agacggggac 3540
 gtggagtgga ccatggcgac gaagtcggat cctggtgggt ttatcccgca atggataact 3600
 20 cgattgagca tccctggagc aatcgcaaaa gatgtgccta gtgtattaaa ctacatacag 3660
 aaataaaaac gtgtcttgat tcattgggtt ggtctctgtt gggttccgag ccaatatttc 3720
 25 acatcatctc ctaaaattct caagaatccc aacgtagcgt agtccagcac gccctctgag 3780
 atcttattta atatcgactt ctcaaccacc ggtggaatcc cgttcagacc attgttaoct 3840
 gtagtgtgtt tgctcttgtt cttgatgaca atgatgtatt tgtcacgata cctgaaataa 3900
 30 taaaacatcc agtcattgag ctattactc gtgaacttat gaaagaactc attcaagccg 3960
 ttcccaaaaa acccagaatt gaagatcttg ctcaactggt catgcaagta gragatcgcc 4020
 35 atgatctgat actttaccaa gctatctctt ccaagttctc ccacgtacgg caagtacggc 4080
 aacgagctct ggaagctttg ttgtttgggg tcata 4115

40 <210> 8
 <211> 3948
 <212> ADN
 <213> *Candida tropicalis*
 45
 <400> 8

gacctgtgac gcttccggtg tcttgccacc agtctccaag ttgaccgacg cccaagtcac 60
 50 gtaccaacttt atttccggtt acacttccaa gatggctggt actgaagaag gtgtcacgga 120
 accacaagct actttctcgg cttgtttcgg tcaaccatct ttggtgttgc acccaatgaa 180
 55 gtacgctcaa caattgtctg acaagatctc gcaacacaag gctaacgcct ggttgttgaa 240
 caccgggttg gtgtgttctt ctgctgctag aggtggtgaa agatgctcat tgaagtacac 300
 60 cagagccatt ttggacgcta tccactctgg tgaattgtcc aagggtgaat acgaaacttt 360
 ccagctcttc aacttgaatg tcccaacctc ctgtccaggt gtcccaagtg aaatcttgaa 420
 cccaaccaag gcttggaccg gaaggtgttg actccttcaa caaggaaatc aagctcttgg 480
 65 ctggttaagtt tgctgaaaac ttcaagacct atgctgacca agctaccgct gaagtgagag 540

ES 2 298 276 T3

ctgcaggctcc agaagcttaa agatatttat tcattattta gtttgccctat ttattttctca 600
 ttacccatca tcattcaaca ctatatataa agttacttgc gatatacattg taatogtgog 660
 5 tgtcgcaatt ggatgatttg gaactgcgct tgaacggat tcatgcacga agcggagata 720
 aaagattacg taatttatct cctgagacaa ttttagccgt gttcadacgc cttcttttgt 780
 10 tctgagcgaa ggataaataa ttagabttcc acagctcat ttaatttccg tcacgcgaat 840
 attgaagggg ggtacatgtg gccgctgaat gtgggggcag taaacgcagt ctctcctctc 900
 ccaggaatag tgcaacggag gaaggataac ggatagaaag cggaatgcga ggaaaatttt 960
 15 gaacgcgcaa gaaaagcaat atccgggcta ccaggttttg agccagggaa cacactccta 1020
 tttctgctca atgactgaac atagaaaaaa caccaagacg caatgaaaag cacatggaca 1080
 20 ttttagacctc cccacatgtg atagtttgc ttaacagaaa agtataataa gaacccatgc 1140
 cgtccctttt ctttcgcgc ttaactttt ttttttttat cttacacaca tcacgacct 1200
 gactgtacac gatattatcg ccacatactt caccaaatgg tacgtgatag taccactogc 1260
 25 tttgattgct tatagagtcc tcgactactt ctatggcaga tacttgatgt acaagcttgg 1320
 tgctaaacca tttttccaga aacagacaga cggctgtttc ggattcaaag ctccgcttga 1380
 30 attgttgaag aagaagagcg acggtaccct catagaotc aactccagc gtatccacga 1440
 tctcgatcgt ccgatatcc caactttcac attcccggtc ttttccatca accttgtcaa 1500
 35 tacccttgag cdygagaaca tcaaggccat cttggccact cagttcaacg atttctcctt 1560
 gggtaaccaga cactcgcact ttgctcctt gttgggtgat ggtatcttta cgttggatgg 1620
 cgccggctgg aagcacagca gatctatgtt gagaccacag tttgccagag aacagatttc 1680
 40 ccacgtcaag ttgttgagc cacacgttca ggtgttcttc aaacacgtca gaaaggcaca 1740
 gggcaagact tttgacatcc aggaattgtt tttcagattg accgtcgact ccgccaccga 1800
 gtttttgttt ggtgaatccg ttgagtctt gagagatgaa tctatcggca tgtccatcaa 1860
 45 tgcgcttgac tttgaaggca aggctggctt tgotgatgt ttaactatt cgcagaatta 1920
 tttggcttgc agagcgggta tgcaacaatt gtactgggtg ttgaacggga aaaagttaa 1980
 50 ggagtgaac gctaaagtgc acaagtttgc tgactactac gtcaacaagg ctttggactt 2040
 gacgcctgaa caattggaaa agcaggatgg ttatgtgtt ttgtacgaat tggtaagca 2100
 55 aaccagagac aagcaagtgt tgagagacca attgttgaac atcatggttg ctggtagaga 2160
 caccaccgcc ggtttgttgt cgtttgtttt ctttgaattg gccagaaacc cagaagttac 2220
 caacaagttg agagaagaaa ttgaggacaa gtttggactc ggtgagaatg ctagtgttga 2280
 60 agacatttcc tttgagtctg tgaagtctg tgaatacttg aaggctgttc tcaacgaac 2340
 cttgagattg taccatccg tgccacagaa tttcagagtt gccaccaaga acactacct 2400
 65

ES 2 298 276 T3

5 cccaagaggt ggtggttaagg acgggttgtc tctgtttttg gtgagaaaag gtcagaccgt 2460
 tatttacggt gtctacgcag cccacagaaa cccagctgtt tacggttaagg acgctcttga 2520
 10 gtttagacca gagagatggt ttgagccaga gacaaagaag cttggctggg ccttctctcc 2580
 attcaacggt ggtccsagaa tctgtttggg acagcagttt gccttgacag aagcttcgta 2640
 tgtcactgtc aggttgcctc aggagtttgc acacttgcct atggaccag acaccgaata 2700
 tccacctaaag aaaatgtcgc atttgaccat gtcgcttttc gacggtgcca atattgagat 2760
 gtattagagg gtcattgtgtc attttgattg tttagtttgc aattactgat taggttaatt 2820
 15 catggattgt tattttattga taggggtttg cgcgtgttgc attcacttgg gatcggtcca 2880
 ggttgatgtt tccctccatc ctgtcgagtc aaaaggagtt ttgttttgta actcgggacg 2940
 atgtttttaa tagaaggctc atctccatgt gattgttttg actgttactg tgattatgta 3000
 20 atctgcggac gttatacaag catgtgattg tggttttgca gccttttgca cgcacaaatga 3060
 tcgtcagacg attacgtaat ctttgtttaga ggggttaaaa aaaacaaat ggcagccaga 3120
 25 atttcaaaaca ttctgcaaac aatgcaaaaa atgggaaact ccaacagaca aaaaaaaaaa 3180
 ctccgcagca ctccgaaccc acagaacaat ggggcgccag aattattgac tattgtgact 3240
 tttttacgct aacgctcatt gcagtgtagt gcgtcttaca cggggtattg ctttctacaa 3300
 30 tgcaggggca cagttgaagg tttgcaacta acgttgcccc gtgtcaactc aatttgaaga 3360
 gtaacttctc aagctcgaat tatgcagctc gtgcgtcaac ctatgtgcag gaaagaaaaa 3420
 atccaaaaaa atcgaaaatg cgactttcga ttttgaataa accaaaaaga aaaatgtcgc 3480
 35 acttttttct cgtctctcgt ctctcgaccc aaatcacaac aaatcctcgc gcgcagtatt 3540
 tcgacgaaac cacaacaaat aaaaaaaca aattctacac cacttctttt tcttcaccag 3600
 40 tcaacaaaaa acaacaaatt atacaccatt tcaacgattt ttgctcttat aaatgctata 3660
 taatggttta attcaactca ggtatgttta ttttactgtt ttcagctcaa gtatgttcaa 3720
 atactaacta cttttgatgt ttgtcgcttt tctagaatca aaacacggcc cacaacacgc 3780
 45 cgagcttgtc gaatagaagg tttgtttact cattagatgg tcccagatta cttttcaagc 3840
 caaagtctct cgagttttgt ttgtgttttc ccaatttct aactatgaag gggttttata 3900
 50 aggtccaaag accccaaggc atagtttttt tggttccttc ttgtcgtg 3948

<210> 9

55 <211> 3755

<212> ADN

<213> *Candida tropicalis*

60

65

ES 2 298 276 T3

<400> 9

	gctcaacaat tgtctgacaa gttctcgcaa cacaaggeta acgcctgggt gttgaacact	60
5	ggttgggttg gttcttctgc tgctagaggt ggtaagagat gttcattgaa gtacaccaga	120
	gccattttgg acgctatcca ctctgggtgaa ttgtccaagg ttgaatacga gactttccca	180
	gtcttcaact tgaatgtccc aacctctcgc ccagggtgac caagtgaat cttgaaccca	240
10	accaaggcct ggaccgaagg tgttgactcc ttcaacaagg aaatcaagtc tttggetggt	300
	aagtttgctg aaaacttcaa gacctatgct gaccaagcta ccgctgaagt tagagctgca	360
15	ggtcagagaag cttaaagata tttattcact atttagtttg cctatttatt tctcatcacc	420
	catcatcatt caacaatata tataaagtta tttcggaact catatatcat tgtaatcgtg	480
20	cgtgttgcaa ttgggttaatt tgaaactgta gttggaacgg attcatgcac gatgcggaga	540
	taacacgaga ttatctctta agacaatttt ggccctcatc acacgcctt cttctgagct	600
	aaggataaat aattagactt cacaagttca ttaaaatatc ogtcacgoga aaactgcaac	660
25	aataaggaag gggggggtag acgtagccga tgaatgtggg gtgccagtaa acgcagtctc	720
	tctctcccc ccccccccc cccctcagg aatagtacaa cgggggaagg ataacggata	780
	gcaagtggaa tgcgaggaaa attttgaatg cgcaaggaaa gcaatatccg ggctatcagg	840
30	ttttgagcca ggggacacac tctcttctg cacaaaaact taacgtagac aaaaaaaaaa	900
	aactccacca agacacaatg aatcgacat ggacatttag acctccccac atgtgaaagc	960
35	ttctctggcg aaagcaaaaa aagtataata aggaccatg ccttccccct tcttggcccg	1020
	tttcaacttt ttctttttct ttgtctatca acacacacac acctcacgac catgactgca	1080
40	caggatatta tgcacacata catcaccaaa tggtagctga tagtaccact cgttttgatt	1140
	gcttataggg tctctgacta cttttacggc agatacttga tgtacaagct tgggtgctaaa	1200
	ccgtttttcc agaaacaaac agacgggtat ttggattca aagctccact tgaattgtta	1260
45	aaaaagaaga gtgaocgtac cctcatagac ttactctcg agcgtatcca agcgtcaat	1320
	cgtecagata tcccaacttt tacattccca atcttttcca tcaaccttat cagcaccctt	1380
50	gagccggaga acatcaaggc tatcttggcc acccagttca acgatttctc cttgggcacc	1440
	agacactcgc actttgctcc tttgttgggc gatggtatct ttaccttga cggtgccggc	1500
	tggaagcaca gcagatctat gttgagacca cagtttgta gagaacagat ttcccacgtc	1560
55	aagttgttg agccacacat gcaggtgttc ttcaagcacg tcagaaaggc acagggcaag	1620
	acttttgaca tccaagaatt gtttttcaga ttgaccgtcg actccgccac tgagtttttg	1680
60	tttgggtgaat ccgttgagtc cttgagagat gaatctattg ggatgtccat caatgcactt	1740
	gactttgacg gcaaggctgg ctttgctgat gcttttaact actcgagaa ctatttggtc	1800
	tcgagagcgg ttatgcaaca attgtactgg gtgttgaacg ggaaaaagtt taaggagtgc	1860
65	aacgctaag tgacacaagt ttgtgactat tacgtcagca agcctttgga cttgacacct	1920

ES 2 298 276 T3

	gaacaattgg aaaagcagga tggttatgtg ttcttgtaag agttggtcaa gcaaaccaga	1980
	gacaggcaag tgttgagaga ccagttgttg aacatcatgg ttgccggtag agacaccacc	2040
5	gccggtttgt tgtcgtttgt tttctttgaa ttggccagaa acccagaggt gaccaacaag	2100
	ttgagagaag aaatcgagga caagtttggg cttggggaag atgctcgtgt tgaagacatt	2160
10	tcctttgagt cgttgaagtc atgtgaatac ttgaaggctg ttctcaacga aactttgaga	2220
	ttgtacccat ccgtgccaca gaatttcaga gttgccacca aaaacactac cttccaagg	2280
	ggaggttgta aggacgggtt atctcctgtt ttggtcagaa aggggtcaac cgttatgtac	2340
15	ggtgtctacg cggccacag aadcccagct gtctacggta aggacgccct tgagtttaga	2400
	ccagagaggt ggtttgagcc agagacaaag aagcttggct gggccttcct tccattcaac	2460
20	ggtggtccaa gaatttgctt gggacagcag ttgccttga cagaagcttc gtatgtcact	2520
	gtcagattgc tccaagagtt tggacacttg tctatggacc ccaacaccga atatccacct	2580
	aggaaaatgt cgcatttgac catgtccctt ttgcaggtg ccaacattga gatgtattag	2640
25	aggatcatgt gttatTTTTG attggtttag tctgtttgta gctattgatt aggttaattc	2700
	acggattgtt atttattgat aggggggtgcg tgtgtgtgtg tgtgttgcat tcacatggga	2760
30	tcgttccagg ttgttgtttc cttccatcct gttgagtcaa aaggagtttt gttttgtaac	2820
	tcgggacgat gtcttagata gaaggtcgat ctccatgtga ttgtttgact gctactctga	2880
	ttatgtaatc tgtaaagcct agacgttatg caagcatgtg attgtggttt ttgcaacctg	2940
35	tttgacgac aaatgatcga cagtcogatta cgtaatcoat attatttga ggggtaataa	3000
	aaaataaatg gcagccagaa ttccaacat ttgcaaaaca atgcaaaaga tgagaaactc	3060
40	caacagaaaa aataaaaaaa ctccgcagca ctccgaacca acaaaacaat ggggggcgcc	3120
	agaattattg actattgtga ctttttttta tttttccgt taactttcat tgcagtgaag	3180
	tgtgttacac ggggttggtg tgggtttggt ttctacaatg caagggcaca gttgaagggt	3240
45	tcacataac gttgcacat atcaactcaa tttatcctca ttcatgtgat aaaagaagag	3300
	ccaaaaggta attggcagac ccccaaggg gaacacggag tagaaagcaa tggaaacacg	3360
50	cccatgacag tgccatttag ccacaaacac atctagtatt cttttttttt ttgtgtgcga	3420
	ggtgcacacc tggactttag ttattgcccc ataaagttaa caatctcacc ttgtgtcttc	3480
	ccagtgtctc cgcctccaga tgctcgtttt acaccctcga gctaacgaca acacaacacc	3540
55	catgagggga atgggcaaag ttaaacactt ttggtttcaa tgattcctat ttgtactct	3600
	cttgttttgt gttttgattt gcaccatgtg aaataaacga caattatata taccttttcg	3660
60	tctgtcctcc aatgtctctt tttgtgcaa tttgtgtttt tgctttttgc ttttgeactc	3720
	tctccactc ccacatcag tgcagcaaca caca	3755

65 <210> 10
<211> 3900

ES 2 298 276 T3

<212> ADN

<213> *Candida tropicalis*

5 <400> 10

	gacatcataa tgaccccggtt atttcgccct caggttgctt atttgagccg taaagtgcag	60
10	tagaaacttt gcttgggtt caaactctag tataatggtg ataactggtt gcactcttgc	120
	cataggcatg aaaataggcc gttatagtac tataattaat aagcgtagga gtataggatg	180
	catatgaccg gttttctctat atttttaaga taatctctag taaattttgt attctcagta	240
15	ggatttcctc aaatttcgca accaattctg gcgaaaaaat gattctttta cgtcaaaagc	300
	tgaatagtgc agtttaagc acctaaaac acatatacag cctctagata cgacagagaa	360
20	gctctttatg atctgaagaa gcattagaat agctactatg agccactatt ggtgtatata	420
	ttagggtattg gtgcaattaa gtacgtacta ataaacagaa gaaaatactt aaccaatttc	480
	tggigtatag ttagtggtga gggacctttt ctgaacattc gggcctaaact tttttttgga	540
25	gtgcgacatc gatttttctg ttgtgtaata atagtgaacc tttgtgtaat aaatcttcat	600
	gcaagacttg cataattcga gcttgggagt tcacgccaat ttgacctcgt tcatgtgata	660
30	aaagaaaagc caaaaggtaa ttagcagacc caatgggaac atggagtga aagcaatgga	720
	agcagcccca ggacggagta atttagtcca cactacatct ggggggtttt tttttgtgcg	780
	caagtacaca cctggacttt agtttttgcc ccataaagtt aacaatctaa cctttggctc	840
35	tccaactctc tccgccccca aatattcgtt ttacacct caagctagcg acagcacaac	900
	acceattaga ggaatggggc aaagttaaac acttttggct tcaatgattc ctattcgcta	960
40	ctacattctt ctcttgtttt gtgctttgaa ttgcaccatg tgaataaac gacaattata	1020
	tatactttt catccctct cctatatctc ttttctctac attttgtttt ttacgtttct	1080
	tgttttgca ctctccct cccacaaaga aaaaaaact acactatgtc gtcttctcca	1140
45	tgtttgccc aagaggttct cgtaccact agtccttaca tcgagtactt tcttgacaac	1200
	tacaccagat ggtactactt catacctttg gtgcttcttt cgttgaactt tataagtttg	1260
	ctccacacaa ggtacttgga acgcaggttc caagccaagc cactcggtaa ctttgtcagg	1320
50	gacctacgt ttggtatcgc tactcgttg ctttctgact acttgaagtc gaaggtacg	1380
	gtcatgaagt ttgcttgggg cctctggaac aacaagtaca tegtcaagaa cccaaagtac	1440
55	aagacaactg ggctcaggat tgttggtctc ccattgattg aaacctgga cccagagAAC	1500
	atcaaggctg ttttggctac tcagttcaat gattctctt tgggaaccag acacgatttc	1560
	ttgtactcct tgttgggtga cggatatttc accttggagc gtgctggctg gaacatagt	1620
60	agaactatgt tgagaccaca gtttctaga gaacaggttt ctcacgtcaa gttgttgag	1680

65

ES 2 298 276 T3

	ccacacggtc aggtgttctt caagcacgtt agaaagcacc gcggtcaaac gttcgacatc	1740
	caagaattgt tcttcagggt gacogtcgac tccgccaccg agttcttgtt tgggtgagtct	1800
5	gctgaatcct tgagggacga atctattgga ttgaccccaa ccaccaagga tttcgatggc	1860
	agaagagatt tcgctgacgc tttcaactat tcgcagactt accaggeeta cagatTTTTg	1920
10	ttgcaacaaa tgtactggat cttgaatggc tcggaattca gaaagtcgat tgctgtcgtg	1980
	cacaagtttg ctgaccacta. tgtgcaaaag gctttggagt tgaccgacga tgacttgacg	2040
	aaacaagacg gctatgtgtt cttgtacgag ttggctaagg aaaccagaga cccaaaggtc	2100
15	ttgagagacc agttattgaa cattttgggt gccggtagag acacgaccgc cggtttgttg	2160
	tcattttgtt tctacgagtt gtcaagaaac cctgaggtgt ttgctaagtt gagagaggag	2220
20	ghggaaaaca gattttggact cgggtgaagaa gctcgtgttg aagagatctc gtttgagtcc	2280
	ttgaagtctt gtgagtactt gaaggctgtc atcaatgaaa octtgagatt gtacccatcg	2340
	gttccacaca acttttagagt tgctaccaga aacactaccc tcccaagagg tgggtgtgaa	2400
25	gatggatact cgccaattgt cgtcaagaag ggtcaagttg tcatgtacac tgttattgct	2460
	accacacagag acccaagtat ctacgggtgcc gacgtgacg tottcagacc agaaagatgg	2520
30	tttgaaccag aaactagaaa gttgggctgg gctaogttc cattcaatgg tggccaaga	2580
	atctgtttgg gtcaacagtt tgccttgacc gaagcttcat acgtcactgt cagattgttc	2640
	caggagtttg cacaottgtc tatggaccca gacaccgaat atocaccaa attgcagaac	2700
35	accttgacct tgtcgtctctt tgatgggtgt gatgttagaa tgtactaagg ttgcttttcc	2760
	ttgctaattt tcttctgtat agcttgtgta tttaaattga atcggcaatt gatttttctg	2820
40	ataccaataa ccgtagtgcg atttgaccaa aacogttcaa actttttgtt ctctcgttga	2880
	cgtgctcgtc catcagcact gtttgaagac gaaagagaaa attttttgta aacaacactg	2940
	tccaaattta cccaacgtga accattatgc aatgagcgg ccttttcaac tggtcgctgg	3000
45	aagcattcgg ggatatctac aacgccotta agtttgaaac agacattgat ttagacacca	3060
	tagatttcag cggcatcaag aatgaccttg cccacatttt gacgaccca acaccactgg	3120
50	aagaatcacg ccagaaacta ggcgatggat ccaagcctgt gaccttgccc aatggagacg	3180
	aagtggagtt gaaccaagcg ttcctagaag ttaccacatt attgtcgaaat gagtttgact	3240
	tggaccaatt gaacgcggca gagtgttat actacgctgg cgacatatcc tacaagaagg	3300
55	gcacatcaat cgcagacagt gccagattgt cttattattt gagagcaaac taatottga	3360
	acatacttgg gtattttgatt togaagcagc gattggattt gatagtcaag gacaacgacg	3420
60	cgttgtttga tagtattttt aaaagttttg aaaagatcta caagttgata agcgtgttga	3480
	acgatatgat tgacaagcaa aaggtgacaa gcgacatcaa cagtctagca ttcacaaatt	3540

65

ES 2 298 276 T3

	gcatcaacta ctgcagaggt caactattct ccgcacacga acttttggga ctggttttgt	3600
	ttggatttgt cgacatctat ttcaaccagt ttggcacatt agacaactac aagaaggta	3660
5	tgccattgat actgaagaac atcagcgatg aagacatctt gatcatacac ttctctccat	3720
	cgacaactaca attgtttaag ctggtgttgg acaagaaga cgacgctgca gttgaacagt	3780
10	tctacaagta catcaacttc acagtgtcac gagactacaa ctccaacatc ggctccacag	3840
	ccaaagatga tategatttg tccaaaacca aactcagtgc ctttgagggtg ttgaccagtt	3900
	<210> 11	
15	<211> 3668	
	<212> ADN	
	<213> <i>Candida tropicalis</i>	
20	<400> 11	
	cctgcagaat tcgcggccgc gtgcacagag tagcagttat gcaagcatgt gatttgtggt	60
25	tttgcaacct gtttgcacga caaatgatcg acagtogatt acgtaatcca tattatttag	120
	aggggtaata aaaaataaat ggcagccaga atttcaaca ttttgcaaac aatgcaaaag	180
30	atgagaaact ccaacagaaa aaataaaaaa actccgcagc actccgaacc aacaaaacaa	240
	tggggggcgc cagaattatt gactattgtg actttttttt attttttcog ttaactttca	300
	ttgcagtga gtgtgttaca cggggtgttg atggtgttgg tttctacaat gcaagggcac	360
35	agttgaagggt ttccacataa cgttgacca tatcaactca atttatctc attcatgtga	420
	taaaagaaga gccaaaaggt aattggcaga ccccccaagg ggaacacgga gtagaaagca	480
40	atggaaacac gcccatgaca gtgccattta gccacaca catctagtat tctttttttt	540
	ttttgtgcgc aggtgcacac ctggacttta gttattgccc cataaagtta acaatctcac	600
	ctttggctct cccagtgtct ccgcctccag atgctcggtt tacaccctcg agctaaagac	660
45	aacacaacac ccatgagggg aatgggcaaa gttaaacact tttggtttca atgattccta	720
	tttgctactc tcttgttttg tgttttgatt tgcaccatgt gaaataaacg acaattatat	780
50	ataccttttc gtctgtctc caatgtctct ttttgctgcc attttgcttt ttgctttttg	840
	cttttgcaact ctctcccaat cccacaactca gtgcagcaac acacaaagaa gaaacatana	900
	aaaacctaca ctatgtcgtc ttctccatcg tttgctcagg aggttctcgc taccactagt	960
55	ccttacatcg agtactttct tgacaactac accagatggg actacttcat ccctttgggtg	1020
	cttctttcgt tgaacttcat cagcttgctc cacacaaagt acttggaaag caggttccac	1080
60	gccaaagccg tcggtaacgt cgtgttggat cctacgtttg gtatcgctac tccgttgatc	1140
	ttgatctact taaagtgcga aggtacagtc atgaagtttg cctggagctt ctggaacaac	1200
65	aagtacattg tcaaagaccc aaagtacaag accactggcc ttagaattgt cggcctccca	1260

ES 2 298 276 T3

	ttgattgaaa ccatagaccc agagaacalc aaagctgtgt tggctactca gttcaacgat	1320
	ttctccttgg gaactagaca cgatttcttg tactccttgt tgggcgatgg tatttttacc	1380
5	ttggacggtg ctggctggaa acacagtaga actatgttga gaccacagtt tgctagagaa	1440
	caggtttccc acgtcaagtt gttggaacca caggttcagg tcttcttcaa qcacgttaga	1500
10	aaacaccgcy gtcagacttt tgacatccaa gaattgttct tcagattgac cgtcgactcc	1560
	gccaccgagt tcttgtttgg tgagictgtt gaatccttga gagacgactc tgttggtttg	1620
	acccaacca ccaaggattt cgaaggdaga ggagatttcg ctgacgcttt caactactcg	1680
15	cagacttacc aggcctacag atttttgttg caacaaatgt actggatttt gaatggcgcg	1740
	gaattcagaa agtcgattgc catcgtgcac aagtttgctg accactatgt gcaaaaggct	1800
20	ttggagttga cggacgatga cttgcagaaa caagacggct atgtgttctt gtaegagttg	1860
	gctaagcaaa cttagagaccc aaaggtcttg agagaccagt tgttgaacat tttggttgcc	1920
	ggttagagaca cgaccgcgcy tttgtgtcg tttgtgttct acgagttgtc gagaaacct	1980
25	gaagtgtttg ccaagttgag agaggaggtg gaaaacagat ttggactcgg cgaagaggct	2040
	cgtgttgaag agatctcttt tgagtccttg aagtcctgtg agtacttgaa ggctgtcatc	2100
30	aatgaagcct tgagattgta cccatctgtt ccacacaact tcagagttgc caccagaaac	2160
	actacccttc caagagggcy tgglaaagac ggatgctcgc caattgttct caagaagggt	2220
	caagttgtca tgtacactgt cattggtagc cacagagacc caagtatcta cggtgccgac	2280
35	gccgacgtct tcagaccaga aagatggttc gagccagaaa ctagaaagtt gggctgggca	2340
	tatgttccat tcaatggtgg tccaagaa:c tgtttgggtc agcagtttgc cttgactgaa	2400
40	gcttcatacg tcaactgtcag attgtccaa gagtttgga acttgtccct ggatccaaac	2460
	gctgagtacc caccaaaatt gcagaacacc ttgacctgt cactctttga tgggtgtgac	2520
	gttagaatgt tctaaggttg cttatccttg ctagtgttat ttatagtttg tgtatttaaa	2580
45	ttgaatcggc gattgatttt tctggtacta ataactgtag tgggttttga ccaaaaccgt	2640
	tcaaaacttt tttttttttt tcttccccct accttcgttg ctcgctcatc agcactgttt	2700
50	gaaaacgaaa aaagaaaatt tttgtaaac aacattgcc aaacttacc aacgtgaacc	2760
	attataacca aatgagcggc gctttcaact ggtcactgga ggcattcggg gatctctaca	2820
	acacccttaa gtttgaggaa gacattgatt tagacaccat agatttcagc ggcataaga	2880
55	atgaccttgt ccacattttg acaaccocaa caccactgga agaatcggc cagaaactag	2940
	gcgatggatc caagcctgtg gccttgccca atggagacga agtggagttg aaccaagcgt	3000
60	tcttagaagt taccacatta ttgtcgacg agtttgactt ggaccaattg aacgcggcgcg	3060
	agttgttata ctaogccggc gacatatcct acaagaaggg cacatcaatt gccgacagtg	3120

65

ES 2 298 276 T3

ccagattgtc ttactatctg agagcaaac acatcttgaa catacttggg tactttatctt. 3180
 cgaagcagcg attggatgtg ataactccg acaacaacgc gttgtttgat aatattttga 3240
 5 aaagtcttga aaagatctac aagttgataa ggcggttgaa cgatatgatt gacaagcaaa 3300
 aggtgacaag cgacatcaac agtetagcat ttatcaactg catcaactac tcgaggggtc 3360
 10 aactattctc cgcacacgaa cttttgggac tggttttgtt tggattgggt gacaactatt 3420
 tcaaccagtt tggctcatta gacaactaca agaaagtatt ggcattgata ctgaagaaca 3480
 tpaagtatga agatatcttg atcgtacgct tctcccatc gacactacaa ttgtttaagc 3540
 15 tgggtgttga taagaaagac gacgccactg ttgaccagtt ctacaagtac atcacctcaa 3600
 cagtgtcgca agctacaaac tccaacatcg gagccacagc caaagatgat atcgatttgt 3660
 ccaaaagcc 3668
 20
 <210> 12
 <211> 3348
 25 <212> ADN
 <213> *Candida tropicalis*
 <400> 12
 30 gatgtgggtgc ttgatttctc gagacacatc cttgtgaggt gccatgaatc tgtacctgtc 60
 tgtaagcaca ggaactgct tcaacacctt attgcatatt ctgtctattg caagcgtgtg 120
 35 ctgcaacgat atctgccaa gttatatgca gaacgtgtcg atggttctct cggtcattatt 180
 ctgttggttag ttctgcaggt aaattttgat gtcaggtagt ggagggaggt ttgtatcgggt 240
 tgtgttttct tcttctctc tctctgatc aaactccacg tctcttctcg gttctgtgtc 300
 40 tgtgtctgag tctactgtt ggattaagtc catcgcatgt gtgaaaaaa gtagecgtta 360
 tttagacaac cagttcgttg ggcgggtatc agaatagtc tgttgtgcac gaccatgagt 420
 45 atgcaacttg acgagacgtc gtttaggaatc cacagaatga tagcaggaag cttactacgt 480
 gagagattct gcttagagga tgttctcttc ttgttgattc cattaggttg gtatcatctc 540
 cgggtggtgac aacttgacac aagcagttcc gagaaccacc cacaacaatc accattccag 600
 50 ctatcaacttc tacatgtcaa cctaegatgt atctcatcac catctagttt cttggcaatc 660
 gtttatttgt tatgggtcaa catccaatac aactccacca atgaagaaga aaaacggaaa 720
 55 gcagaatacc agaattgacg tgtgagttcc tgaccattgc taatctatgg ctatatctag 780
 tttgtctatcg tgggatgtga tctgtgtgtc cttcatttgc gtttgtgttt atttcgggta 840
 tgaatattgt tatactaaat acttgatgca caaacatggc gtcgagaaa tcgagaatgt 900
 60 gatcaacgat gggttctttg ggttccgctt acctttgcta ctcatgagag ccagcaatga 960
 gggccgactt atcgagttca gtgtcaagag attcgagtcg gcgccacatc cacagaacaa 1020
 65 gacattggtc aaccgggcat tgagcgttcc tgtgatactc accaaggacc cagtgaatat 1080

ES 2 298 276 T3

caaagcgatg ctatcgaccc agtttgatga cttttccctt gggttgagac tacaccagtt 1140
 tgcgcggttg ttggggaaaag gcattctttac tttggacggc ccagagtggg agcagagccg 1200
 5 atctatgttg cgtccgcaat ttgccaaaaga tcgggtttct catatcctgg atctagaacc 1260
 gcattttgpg ttgcttcgga agcacattga tggccacaat ggagactact tcgacatcca 1320
 10 ggagctctac ttccggttct cgatggatgt ggcgacgggg ttttctgttg gcgagtctgt 1380
 ggggtcggtg aaagacgaaag atgcgaggtt cctggaagca ttcaatgagt cgcagaagta 1440
 tttggcaact agggcaacgt tgcacgaggt gtactttctt tgtgacgggt ttaggtttcg 1500
 15 ccagtlacaac aaggttgtgc gaaagttctg cagccagtgt gtccacaagg cgtagatgt 1560
 tgcaccgga gacaccagcg agtacgtgtt tctccgcgag ttggtcaaac aactcggaga 1620
 20 tccggttgtt ttacaagacc aagcgttgaa cgtcttgctt gctggacggg acaccaccgc 1680
 gtggttatta tcgtttgcaa catctgagct agcccggaat gaccacatgt ggaggaagct 1740
 acgagaggag gttatcctga cgatgggacc gtccagtgtt gaaataaccg tggccgggtt 1800
 25 gaagagttgc cgttacctca aagcaatcct aaacgaaact cttcgactat acccaagtgt 1860
 gcctaggaac gcgagatttg ctacgaggaa tacgacgctt cctcgtggcg gaggtccaga 1920
 30 tggatcgttt ccgattttga taagaaaggg ccagccagtg gggatattca tttgtgctac 1980
 aacttgaat gagaaggtat atgggaatga tagccatgtg tttcgaccgg agagatgggc 2040
 tgcgttagag ggtaagagtt tgggctggtc gtatcttcca ttcaacggcg gcccgagaag 2100
 35 ctgccttggg cagcagtttg caatccttga agcttcgtat gttttggctc gattgacaca 2160
 gtgctacacg acgatacagc ttagaactac cgagtaccca ccaagaaac tcgttcatct 2220
 40 caogatgagt cttctcaacg ggggtgtacat ccgaactaga acttgattat gtgtttatgg 2280
 ttaatogggg caaagcactg caagtcattg atgtttgtgg aagccagca ttggtgttcc 2340
 ggagcatcaa taaccaatgt cttgaagggt ttgattttct tgaccttctt cttcctgagc 2400
 45 ttctttccgt caaacttgta cagaatggcc atcatttcag gaacaaccac gtacgacggc 2460
 cggtagcgca tctggagtat ctgcgcgtcg ttcaagtagc acgaaaacag caacgacgtc 2520
 50 accatctgct tcccaatctt gacacccccc gataccctg cggcttcctg gatcaaaaac 2580
 gtcggaacc ccgcgtatat gtccatgtaa ttctccatgg ccacctccat caacacactg 2640
 55 atggagcgac tgacgggtgc accactgcc cgggttgagt caaggcagta tgatgcggg 2700
 atccagtact ccaatgggaa cctctgcacg gtgtcgtgc agtttttgag gcgtatttcg 2760
 atccatgac gttctttggt gctgtagtat aacgagctct tgggtgcctt gaaatggaac 2820
 60 aggttgatg tgttgttgag tttgtctgcg tgcttggtt gcaagtctc gatcgagcgt 2880
 agtgagtaga cagttggcgg ggggtggggc tcgggcttta ttctgtgtt gtgtttcctt 2940

65

ES 2 298 276 T3

	cttagtcttg gaatgacgct gttatcgacg gttcgtagta taagtagcgc caatatgaga	3000
	atgtatatcc gcatcaccca agactcttca gcoctgttaca accactgagg ctgttcggccg	3060
5	tgtgaccaat tggtttcttt ggtgacctag attggtcccg cagggaagc aagggttgct	3120
	aggggggcat accaaacaag gtctgtgaat cagtatctat ggtgctacca tctgtgtggt	3180
10	tggggggaaa ttcccgcat tttgtgtaac gaaagttcta gaaagttctc gtgggttctg	3240
	agaatctgct ggaacctatc acccgcatctt ccggtgccaa agtgggaaga gcaatcaacc	3300
	cacctgtgct tgcaccaatca gccattcccc tgggaatata aattcaac	3348
15	<210> 13	
	<211> 3826	
	<212> ADN	
20	<213> <i>Candida tropicalis</i>	
	<400> 13	
25	tggagtcgcc agacttgctc adttttgact ccccttcgaaa ctcaaagtac gttcaggcgg	60
	tgtctaacga aacgctccgt atctaccggg gggtagccag aaacatgaag acagctacgt	120
30	gcaacacgac gttgccacgc ggaggaggca aagacggcaa ggaacctatc ttggtgcaga	180
	agggacagtc cgttgggttg attactattg ccacgcagac ggacccagag tattttgggg	240
	ccgacgctgg tgagtttaag ccggagagat ggtttgattc aagcatgaag aacttggggt	300
35	gtaaataactt gccgttcaat gctgggccac ggacttgctt ggggcagcag tacactttga	360
	ttgaagcgaq ctacttgcta gtccggttgg ccagacctc ccgggcaata gatttgcagc	420
40	caggatcggc gtacccacca agaaagaagt cgttgatcaa catgagtgct gccgacgggg	480
	tgtttgtaaa gctttataag gatgtaacgg tagatggata gttgtgtagg aggagcggag	540
	ataaattaga ttgatatttg tgtaaggttt tggatgtcaa cctactccgc acttcatgca	600
45	gtgtgtgtga cacaaggggt tactacgtgt gcgtgtgcgc caagagacag cccaaggggg	660
	tggtagtggt tgttggcgga agtgcgtgtg acacaacggc tgggttcttg ccaatggtgg	720
50	actaagtgca ggtaagcagc gacctgaaac attcctcaac gcttaagaca ctggtggtag	780
	agatgoggac caggtctattc ttgtcgtgct acccggcgca tggaaaatca actgcgggaa	840
	gaataaattt atccgtagaa tccacagagc ggataaattt gccacctcc atcatcaacc	900
55	acgcggccac taactacatc actccctat tttctctctc tctctttgtc ttactccgt	960
	cccgtttctt tagccacaga tacacaccca ctgcaaacag cagcaacaat tataaagata	1020
60	cgccaggccc accttcttct tttttcttca cttttttgac tgcaactttc tacaatccac	1080
	cacagccacc accacagcgg ctatgattga acaactccta gaatattggt atgtcgttgt	1140
65	gccagtgttg tacatcatca aacaactcct tgcatacaca aagactcgcg tcttgatgaa	1200

ES 2 298 276 T3

	aaagttgggt gctgctccag tcacaaacaa gttgtacgac aacgctttcg gtatcgtcaa	1260
	tggatggaag gctctccagt tcaagaaaga gggcagggct caagagtaca acgattacaa	1320
5	gtttgaaccac tccaagaacc caagcgtggg cacctacgtc agtattcttt tcggcaccag	1380
	gatcgtcgtg accaaagatc cagagaatat caaagctatt ttggcaaccc agtttggtga	1440
10	ttttcttttg ggcaagaggc acactctttt taagcctttg ttaggtgatg ggatcttcac	1500
	attggacggc gaaggctgga agcacagcag agccatgttg agaccacagt ttgccagaga	1560
	acaagttgct catgtgacgt cgttgggacc acacttccag ttgttgaaga agcatattct	1620
15	taagcacaag ggtgaatact ttgatatcca ggaattgttc tttagattta ccgttgatc	1680
	ggccscggag ttcttatttg gtgagtcctg gcactcctta aaggacgaat ctattggtat	1740
20	caaccaagac gatatagatt ttgctggtag aaaggacttt gctgagtcgt tcaacaaagc	1800
	ccaggaatac ttggctatta gaaccttggg gcagacgttc tactggttgg tcaacaacaa	1860
	ggagittaga gactgtacca agctggtgca caagttcacc aactactatg ttcagaaagc	1920
25	tttgatgct agcccagaag agcttgaaaa gcaaagtggg tatgtgttct tgtacgagct	1980
	tgtcaagcag acaagagacc ccaatgtgtt gcgtgccag tctttgaaca tcttgttggc	2040
30	cggaagagac accactgctg ggttgttgtc gtttgcgtgc tttgagttgg ccagacaccc	2100
	agagatctgg gccaaagtga gagaggaaat tgaacaacag ttgggtcttg gagaagactc	2160
	tcgtgttgaa gagattacct ttgagagctt gaagagatgt gagtacttga aagcgttctt	2220
35	taatgaaacc ttgcgtatctt acccaagtgt cccaagaaac ttcagaatcg ccaccaagaa	2280
	cacgacattg ccaagggggc gtggttcaga cggtagctcg ccaatcttga tccaaaaggg	2340
40	agaagctgtg tcgtatggtg tcaactctac tcatttggac cctgtctatt acggccctga	2400
	tgctgctgag ttcagaccag agagatggtt tgagccatca accaaaaagc tcggctgggc	2460
	ttacttgcca ttcaacgggtg gtccaagaat ctgtttgggt cagcagtttg ccttgacgga	2520
45	agctggctat gtgttggtta gattggtgca agagttctcc cacgttaggc tggacccaga	2580
	cgaggtgtac ccgccaagaa ggttgaccac cttgaccatg tgtttgcagg atggtgctat	2640
50	tgtcaagttt gactagcggc gtggtgaatg cgtttgattt tgtagtttct gtttgcagta	2700
	atgagataac tattcagata aggcgagtgg atgtacgttt tgtaagagtt tccttacaac	2760
	cttgggtggg tgtgtgaggt tgaggttgca tcttggggag attacacctt ttgcagctct	2820
55	ccgtatacac ttgtactctt tgtaacctct atcaatcatg tggggggggg ggttcattgt	2880
	ttggccatgg tgggtgatgt taaatccgcc aactacccaa tctcacatga aactcaagca	2940
60	cactaaaaaa aaaaaagatg ttgggggaaa actttggttt cccttcttag taattaaaca	3000
	ctctcactct cactctcact ctctccactc agacaaacca accacctggg ctgcagacaa	3060

65

ES 2 298 276 T3

ccagaaaaaa aaagaacaaa atccagatag aaaaacaaag ggcgtggacaa ccataaataa 3120
 acaatctagg gtctactcca tcttcacttg tttcttcttc ttcagactta gctaacaaac 3180
 5 aactcacttc accatggatt acgcaggcat cagcgtggc tccatcagag gcgaggcctt 3240
 gaagaaactc gcagaattga ccatccagaa ccagccatcc agcttgaaag aaatcaaac 3300
 10 cggcatccag aaggacgact ttgccaagtt gttgtctgcc accccgaaaa tccccaccaa 3360
 gcacaagttg aacggcaacc acqaattgtc tgaggtcgcc attgccaaaa aggagtaoga 3420
 ggtgttgatt gccttgagcg accccacaaa agacccaatc aaagtgacct ccagatcaa 3480
 15 gatcttgatt gacaagttaa aggtgtactt gtttgagttg cctgaccaga agttctctta 3540
 ctccatcgtg tccaactccg tcaacatcgc cccctggacc ttgctcgggg agaagttgac 3600
 caccggcttg atcaacttg ccttcacaga caacaagcag cacttgagcg aggtcattga 3660
 20 catcttcaac gagttcatcg accagttctt tggcaacacg gagccgcaat tgaccacatt 3720
 cttgaccttg tgcggtgtgt tggacgggtt gattgacctt gccaaactct tgagcgtgtc 3780
 25 ctgcgggacc ttcaagatct tcttgaactt ggactcgtat gtggac 3826

<210> 14

<211> 3910

30 <212> ADN

<213> *Candida tropicalis*

<400> 14

35 ttacaatcat ggagctcgtt aggaaccacg atgtctggga gaagctccgc gaagaggtca 60
 aacgaactt tggcatggag tcgccagact tgctcacttt tgactctctt agaagctcaa 120
 40 agtacgttca ggcggtgctc aacgaaacgc ttcgtatcta cccgggggtg ccacgaacaa 180
 tgaagacagc tacgtgcaac acgacgttgc cgcgtggagg aggcaaagac ggtaaggaaac 240
 ctatttttggg gcagaagggc cagtcggttg ggttgattac tattgccacg cagacggacc 300
 45 cagagtattt tggggcagat gctggtgagt tcaaaccgga gagatggttt gattcaagca 360
 tgaagaactt ggggtgtaag tacttgccgt tcaatgctgg gccccggact tgtttggggc 420
 50 agcagtaaac ttgattgaa gcgagctatt tgctagtcag gttggcgcag acctaccggg 480
 taatcgattt gctgccaggg tcggcgtaac caccaagaaa gaagtcgttg atcaatatga 540
 55 gtgctgccga tggggtggtt gtaaagtttc acaaggatct agatggatat gtaagggtgtg 600
 taggggggac ggagataaat tagatttget tttgtgtaag gtttagcagc tcaagotact 660
 ccgcactttg tgtgtaggga gcacatactc cgtctgcgcc tgtgccaaaga gacgggccag 720
 60 gggtagtgtg tgggtggtga agtgcatgtg acacaatacc ctggttcttg ccaattgggg 780
 atttagtgtg ggtaagctgc gacctgaaac actcctcaac gcttgagaca ctggtgggta 840
 65 gagatgcggg ccaggaggct atttttgtcg tgctaccctt gcacggaaaa tcgattgagg 900

ES 2 298 276 T3

gaagaacaaa tttatccgtg aaatccacag agcggataaa tttgtcacat tgctgcgttg 960
 cccacccaca gcattctctt ttctctctct ttgtcttact ccgtctctgt ttccttatcc 1020
 5 agaaatacac accaactcat ataaagatac gctagcccag ctgtctttct ttttcttcac 1080
 ttttttttgt gtgttgettt ttgtgtgtgt actttctaca accaccacca ccaccaccac 1140
 10 catgattgaa caaatcctag aatattggta tattgttgtg cctgtgttgt acatcatcaa 1200
 acaactcatt gcctacagca agactcgcgt ctgtatgaaa cagttgggtg ctgtccaat 1260
 cacaaaccag ttgtacgaca acgttttcgg tatcgtcaac ggatggaaag ctctccagtt 1320
 15 caagaaagag ggcagagctc aagagtacaa cgatcacaaag ttgtacagct ccaagaaccc 1380
 aagcgtcggc acctatgtca gtattctttt tggcaccaag attgtcgtga ccaaggatcc 1440
 20 agagaatatc aaagctattt tggcaaccca gtttggcgat ttttctttgg gcaagagaca 1500
 cgctcttttt aaacctttgt tagtgtgatgg gatcttcacc ttggacggcg aaggctggaa 1560
 gcatagcaga tccatgttaa gaccacagtt tgccagagaa caagttgctc atgtgacgtc 1620
 25 gttggaacca cacttcagtt tgttgaagaa gcatactctt aaacacaagg gtgagtactt 1680
 tgatatccag gaattgttct ttagatttac tgtcgaactc gccacggagt tcttatttgg 1740
 30 tgagtcogtg cactccttaa aggaagaaac tatcgggtac aaccaagacg atatagattt 1800
 tgctggtaga aaggactttg ctgagtcgtt caacaaagcc caggagtatt tgtctattag 1860
 aatttttgtg cagaccttct actggttgat caacaacaag gagtttagag actgtaccaa 1920
 35 gctggtgcac aagtttacca actactatgt tcagaaagct ttggatgcta cccagagga 1980
 acttgaagag caaggcgggt atgtgttctt gtatgagctt gtcaagcaga cgagagaccc 2040
 40 caaggtgttg cgtgaccagt ctttgaacat cttgttggca ggaagagaca ccactgctgg 2100
 gttgttgtcc tttgctgtgt ttgagttggc cagaaaccca cacatctggg ccaagttgag 2160
 agaggaaatt gaacagcagt ttggtcttgg agaagactct cgtgttgaag agattacctt 2220
 45 tgagagcttg aagagatgtg agtacttgaa agcgttccct aacgaaacct tgcgtgttta 2280
 cccaagtgtc ccaagaaact tcagaatcgc caccaagaat acaacattgc caaggggtgg 2340
 50 tgggtccagac ggtaccagc caatcttgat ccaaaagga gaaggtgtgt cgtatggtat 2400
 caactctacc cacttagatc ctgtctatta tggcctgat gctgctgagt tcagaccaga 2460
 gagatggttt gagccatcaa ccagaaagct cggctgggct tacttgccat tcaacgggtg 2520
 55 gccacgaatc tgtttgggtc agcagtttgc cttgacogaa gctggttacg ttttggtcag 2580
 attggtgcaa gagttctccc acattaggct ggacccagat gaagtgtatc caccaaagag 2640
 60 gttgaccaac ttgaccatgt gtttgcagga tgggtctatt gtcaagtttg actagtacgt 2700
 atgagtgctt ttgattttgt agtttctgtt tgcagtaatg agataactat tcagataagg 2760

ES 2 298 276 T3

5 cgggtggatg tacgttttgt aagagtttcc ttacaaccct ggtgggtgtg tgaggttgca 2820
 tottagggag agatagcacc ttttcagct ctccgtatac agttttactc ttgttaacct 2880
 atgecaatca tgtggggatt cattgtttgc ccattggtgt gcatgcaaaa tccccccaac 2940
 tacccaatct cacatgaaac tcaagcacac tagaaaaaaa agatgtttgcg tgggttcttt 3000
 10 tgatgttggg gaaaactttc gtttctttc tcaglaatta aacgttctca ctccagcaaa 3060
 ccacotgggc tgcagacaac cagaaaaaac .aaaatccaga tagaagaaga aaggcctgga 3120
 caaccataaa taacaacct aggtccact ccattcttca cttcttcttc ttcagactta 3180
 15 tctacaaac gattcacttc accatggatt acgcaggtat caagcgtggg tccatcagag 3240
 gcgaagcctt gaagaaactc gccgagttga ccattccagaa ccagccatcc agcttgaaag 3300
 aaatcaacac dggcatccag aaggacgact ttgccaaagt gttgtcttcc accccgaaaa 3360
 20 tccacaccaa gcacaagttg aatggcaacc acgaattgtc cgaagtgcgc attgccaaaa 3420
 aggagtacga ggtgttgatt gccttgagcg acgccacgaa agaaccaatc aaagtccact 3480
 cccagatcaa gatcttgatt gacaagtcca aggtgtactt gtttgagttg ccgaccaga 3540
 25 agttctctca ctccatcgtg tccaaactcc ttaacattgc cccctggacc ttgctcgttg 3600
 agaagttgac caccggcctg atcaacttgg cgttccagaa caacaagcag cacttggacg 3660
 30 aagtcacga catcttcaac gagttcatcg acaagttctt tggcaacaca gagccgcaat 3720
 tgaccaactt cttagacctg tccggtgtgt tggacgggtt gattgaccat gccaaactct 3780
 tgagcgtgtc ctccaggacc ttcaagatct tcttgaactt ggactcgttt gtggacaact 3840
 35 cggacttctt gaaccagctg gagaactact ccgacttttt gtacgacgag ccgaacgagt 3900
 accagaactt 3910

40 <210> 15
 <211> 3150
 <212> ADN
 45 <213> *Candida tropicalis*
 <400> 15

50 gaattctttg gatctaattc cagctgatct tgctaactct tatcaacgta gtttgtatca 60
 ttgtttgtct gaattataca caccagtggg agaatatggt ctaatttga cgtcccaactg 120
 gcatttgttg tttgtggggg ggggggggtg cacacatttt tagtgccatt ctttgttgat 180
 55 taccctctcc ccctatcatt cattcccaaa ggattagttt tttctcact ggaattcgct 240
 gtccaaactgt caaccccccc cccccccccc ccaactgccc taccctgccc tgccctgcac 300
 gtctctgtgt ttgtgtgtgt ttttcccaac gctataaaag ccttggcgtc cggccaaagt 360
 60 ttttccaccc agccaaaaaa acagtctaaa aaatttggtt gatccttttt ggttgcaagg 420

65

ES 2 298 276 T3

ttttccacca ccacttccac cacttcaact attcgaacaa aagatgctcg atcagatctt 480
 acattactgg tacattgtct tgccattggt gccattatc aaccagatcg tggctcatgt 540
 5 caggaccaat tatttgatga agaaattggg tgctaagcca ttcacacacg tccaacgtga 600
 cgggtggttg ggcttcaaat tcggccgtga attcctcaaa gcaaaaagtg ctgggagact 660
 10 ggttgattta atcatctccc gtttccacga taatgaggac aotttotoca getatgcttt 720
 tggcaaccat gtggtgttca ccagggaccc cgagaatatc aaggcgcttt tggcaaccca 780
 gtttggtgat ttttcatttg gcagcadggg caagttcttc aaaccattat tggggtacgg 840
 15 tatcttcaca ttggacgccg aadgctggaa gcacagcaga gccatgttga gaccacagtt 900
 tgccagagaa caagttgctc atgtgacgtc gttggaacca cacttccagt tgttgaagaa 960
 20 gcatatcctt aaacacaagg gtgagtactt tgatatccag gaattgttct ttagatttac 1020
 tgtcgactcg gccacggagt tcttatttgg tgagtccgtg cactccttaa aggacgagga 1080
 aattggctac gacacgaaag acatgtctga agaaagacgc agatttgccg acgcgttcaa 1140
 25 caagtcgcaa gtctacgtgg ccaccagagt tgctttacag aacttgtaact ggttggtcaa 1200
 caacaaagag ttcaaggagt gcaatgacat tgtccacaag ttaccaact actatgttca 1260
 30 gaaagccttg gatgetacco cagaggaact tgaaaagcaa ggcgggtatg tgttcttgta 1320
 tgagcttgtc aagcagacga gagaccccaa ggtgttgogt gaccagtctt tgaacatctt 1380
 gttggcagga agagacacca ctgctgggtt gttgtccttt gctgtgtttg agttggccag 1440
 35 aaaccacac atctgggoca agttgagaga ggaaattgaa cagcagtttg gtcttgagaa 1500
 agactctcgt gttgaagaga ttacctttga gagcttgaag agatgtgagt acttgaaggc 1560
 40 cgtgttgaac gaaactttga gattacaccc aagtgtccca agaaacgcaa gatttgccat 1620
 taaagacacg actttaccaaa gaggcgggtg ccccaacggc aaggatccta tcttgatcag 1680
 gaaggatgag gtggtgcagt actccatctc ggcaactcag acaaactctg cttattatgg 1740
 45 cgccgatgct gctgatttta gacgggaaag atggtttgaa ccatcaacta gaaacttggg 1800
 atgggctttc ttgccattca acggtggtcc aagaatctgt ttgggacaac agtttgcttt 1860
 50 gactgaagcc ggttacgttt tgggttagac tgttcaggag ttccaaaact tgtcacaga 1920
 cccogaaacc aagtaccac cacttagatt ggcacacttg acgatgtgct tgtttgacgg 1980
 tgcacacgtc aagatgtcat aggttctccc atacaagtag ttcagtaatt atacactgtt 2040
 55 tttactttct ctcatacca aatggacaaa agttttaagc atgcctaaca acgtgaccgg 2100
 acaatttgtt cgcactagta tgaacaatt gtaaaaatag tgtacactaa tttgtggtgg 2160
 60 ccggagataa attacagttt ggttttgtgt aaactcggcg atatctctgg cagtttctct 2220
 tctccgcagc agctttgcca cgggtttgct ctggggccaa caaattcaaa agggggagaa 2280

65

ES 2 298 276 T3

acttaacacc ccttatctct ccaactctagg ttgtagctct tgtggggatg caattgtcgt 2340
 acgtttttta tgttttgtct agactttgat gattacgttg gatitcttat gtctgaggcg 2400
 5 tgcctgaaag aagtgctcaa atgtgacagg cgacgctatt cgacatgaac gcgaaaggggt 2460
 tatttgcate aatacgaggg gctgactcta gtctaggatg gcagtcctag gttgcaaaca 2520
 tgttgcaacca tatccctcct ggagttgggc gccctcgccct acgccaccct cagcgatcgg 2580
 10 cactttccgt tgttcaatat ttctccttcc cattgttcca ggggttatca acaacgttgc 2640
 cggcctcctc tccaaattac aagaaaaata aattgtcgca cggcaccgat ctgtcaaaga 2700
 tacagataaa ccttaaatct gcaaaaacaa gacccctccc catagcctag aagcaccagc 2760
 15 aagatgatgg agaaactcct ccagtactgg tacatcgac tctctgtatg gttcatcctt 2820
 cgctacttgg ctcccacgc acgagccgtc taettgcgc acaagctcgg cgcggcgcca 2880
 20 ttcagcaca cccagtcga cggctggtat ggggttaagt ttggggggga gtttctcaag 2940
 gcgaagaaga tcggcgcgca gacggacttg gtgcatgcgc ggttccgtgg cggcatggac 3000
 accttctcga gctacacttt cggcatccat atcatcctta cccgggaacc ggagaacatc 3060
 25 aaggcggtct tggcgacgca gttcgatgac ttctcgctcg gtggcaggat caggttcttg 3120
 aagcogttgt tgggctatgg gatattcacg 3150'

30 <210> 16
 <211> 3579
 <212> DNA
 <213> *Candida tropicalis*
 35
 <400> 16

40 aaaaccgata caagaagaag acagtcaaca agaacgttaa tgtcaaccag gcgccaaaga 60
 gacggttttg cggacttga agaattgtggc atttgcccat gatgtttatg ttctggagag 120
 gtttttcaag gaatcgtcat cctccgccac cacaagaacc accagttaac gagatccata 180
 45 ttcaacaacc acogcaaggt gacaatgctc aacaacaaca gcaacaacaa caacccccac 240
 aagaacagtg gaataatgcc agtcaacaaa gagtggtgac agacgagggg gaaaacgcaa 300
 gcaacagtgg ttctgatgca agatcagcta caccgcttca tcaggaaaag caggagctcc 360
 50 caccaccata tgcccatcac ggcgaacacc agcaggttag tgtatagtag tctgtagtta 420
 agtcaatgca atgtaccaat aagactatcc cttottacaa ccaagttttc tgccgcgcct 480
 55 gtctggcaac agatgctggc cgacacactt tcaactgagt ttggtctaga attcttgcac 540
 atgcacgaca aggaaactct tacaaagaca acacttgtgc tctgatgcca cttgatcttg 600
 ctaagcctta tcaacgtaat tgagatcatt gtttgtctga attatacaca ccagtggag 660
 60 aatctggtct aatctgcacg cctcatgggc attgtgtgtt ttgggggggg gggggggggt 720
 gcacacattt ttagtgcgaa tgtttgtttg ctggttcccc ctccccctc cccoctatca 780

65

ES 2 298 276 T3

	tgccacaggg attagttttt tctcactgg aattcgctgt ccacctgtca accccctcac	840
	tgccctgccc tgccctgcac gccctgtgtt ttgtgctgtg gcactccac gctataaaag	900
5	ccctggcgta cggccaaggt ttttctcac agccaaaaaa aaatttggt gatccttttg	960
	ggctgcaagg tttttacca ccaccaccac caccacctca actattbaaa caaaggatgc	1020
10	tcgaccagat cttccattac tggtaattg tcttgccatt gttggtcatt atcaagcaga	1080
	tcgtggctca tgccaggacc aattatttga tgaagaagt gggcgctaag ccattcacac	1140
	atgtccact agacgggtgg ttggcttoa aatttgccg tgaattctc aaagctaaaa	1200
15	gtgtggggag gcaggttgat ttaatcatct cccgtttcca cgataatgag gacactttct	1260
	ccagctatgc ttttggaac catgtggtgt tcaccaggga ccccgagaat atcaaggcgc	1320
20	ttttggcaac ccagtttggg gatttttcat tgggaagcag ggtcaaattc ttcaaaccat	1380
	tggtggggta cggtatcttc accttgagc gcgaaggctg gaagcacagc agagccatgt	1440
	tgagaccaca gtttgccaga gagcaagttg ctcattgtac gtcttgga ccacatttcc	1500
25	agttgtttaa gaagcatatt ctttaagcaca aggggtgaata ctttgatata caggaattgt	1560
	tctttagatt taccgttgat tcagcgacgg agttcttatt tgggtgagtcc gtgcactcct	1620
30	taagggacga ggaattggc tacgatacga aggacatggc tgaagaaaga cgcaaatttg	1680
	ccgacgcgtt caacaagtcg caagtctatt tgtccaccag agttgcttta cagacattgt	1740
	actggttggg caacacaaaa gagttcaagg agtgcaacga cattgtccac aagttcacca	1800
35	actactatgt tcagaaagcc ttggatgcta cccagaggga acttgaaaaa caaggcgggt	1860
	atgtgttctt gtacgagctt gccaaacaga cgaagaccc caatgtgttg cgtgaccagt	1920
40	ctttgaacat cttgttggct ggaagggaca ccactgctgg gttgttgtcc tttgctgtgt	1980
	ttgagttggc caggaaccca cacatctggg ccaagttgag agaggaaatt gaatcacact	2040
	ttgggctggg tgaggactct cgtgttgaag agattacctt tgagagcttg aagagatgtg	2100
45	agtactttaa agccgtgttg aacgaaacgt tgagattaca cccaagtgtc ccaagaaacg	2160
	caagatttgc gattaaagac acgactttac caagaggcgg tggccccaac ggcaaggatc	2220
50	ctatcttgat cagaaagaat gaggtggtgc aatactccat ctgggcaact cagacaaate	2280
	ctgcttatta tggcgccgat gctgctgatt ttagaccgga aagatggttt gagccatcaa	2340
	ctagaaactt gggatgggct tacttgccat tcaacgggtg tocaagaatc tgcttgggac	2400
55	aacagtttgc tttgacogaa gccggttacg ttttggttag acttggtcag gaattcccta	2460
	gcttgtcaca ggaccccgaa actgagtacc caccacctag attggcacac ttgacgatgt	2520
60	gcttgttga cggggcatac gtcaagatgc aataggtttt gggttgactt tgtttcata	2580
	tgcaagtagt tcagtaatta cacactaatt tgtggtggcc ggcgataaat taccgtttgg	2640

65

ES 2 298 276 T3

5 ttttgtgtaa aaattcggac atctctggtg gtttcccttc tccgcagcag ctttgccacg 2700
 ggttigctct ggggccaaca aattcgaaag gggggggggg gggggagaaa gttaacaccc 2760
 cctgttccca ccgtaggctg tagctcttgt ggggggatgt aattgtcgta cgttttcatg 2820
 tttggcccag actttgatga ttacgtaggc ttctttatgt ctaaggcgtg cttgacacaa 2880
 10 gtgtcaaaag gtgacaggcg acgttattcg acatgaacgc aaaagggtta tttgcacgca 2940
 taagaggggt tgctctggt ctaagaagga ccccccaggt tgaaacatg ttgcaactgca 3000
 tcccaactcag agtttggtga ccacgcctac gctaacctc agcgatcggc actttccgtt 3060
 15 gctcaatatt tctctccccc ctgcttcccc ccattgttcc agggattatc aacaacgttg 3120
 ctggctctct cttccccccc tccccccagt tatgtacaag aaaattaat tgtgcacagg 3180
 cacogatacg tbaaagatac agagaaacct taatccctcc catagcctag aagcatcaaa 3240
 20 aagatgattg agcaactcct ccagtactgg tacattgcac tccctgtatg gttcattctc 3300
 cgctacgtgg cttcccacgc acgaaccatc tacttgccgc acaagctcgg cgcggcgccg 3360
 25 ttcacgcaca cccagtacga cggatggtat ggggtcaggt ttggggggga gtttctcaag 3420
 gcaagaaga ttggaaggca gacggacttg gtgcacgcgc ggttccgttg agggggcatg 3480
 gatactttct cgagctatac ttccggcatc catatcattc ttactcggga cccggagaac 3540
 30 atcaaggcgg tcttggcgac gcagttcgat gacttttcg 3579

<210> 17
 35 <211> 31
 <212> ADN
 <213> cebador
 40 <400> 17

ccttaattaa gaggtcgtg gttgagttt c

31

45 <210> 18
 <211> 29
 <212> ADN
 50 <213> cebador
 <400> 18

ccttaattaa ttgataatga cgttgcggg

29

<210> 19
 <211> 33
 60 <212> ADN
 <213> cebador
 <400> 19

aggcgcgccg gagtccaaa agaccaacct ctg

33

ES 2 298 276 T3

	<210> 20	
	<211> 34	
	<212> ADN	
5	<213> cebador	
	<400> 20	
10	ccttaattaa tacgtggata ccttcaagca agtg	34
	<210> 21	
	<211> 35	
15	<212> ADN	
	<213> cebador	
	<400> 21	
20	ccttaattaa gctcacgagt ttggtgattt tcgag	35
	<210> 22	
25	<211> 35	
	<212> ADN	
	<213> cebador	
30	<400> 22	
	gggtttaaac cgcagagggtt ggtcttttg gactc	35
35	<210> 23	
	<211> 10	
	<212> ADN	
40	<213> cebador	
	<400> 23	
45	gggtttaaac	10
	<210> 24	
	<211> 9	
50	<212> ADN	
	<213> cebador	
	<400> 24	
55	aggcgegcc	9
	<210> 25	
60	<211> 10	
	<212> ADN	
	<213> cebador	
65	<400> 25	
	ccttaattaa	10

ES 2 298 276 T3

	<210> 26	
	<211> 21	
	<212> ADN	
5	<213> cebador	
	<400> 26	
10	tcycaaacwg gtacwgcwga a	21
	<210> 27	
	<211> 21	
15	<212> ADN	
	<213> cebador	
	<400> 27	
20	ggtttgggta aytcwactta t	21
	<210> 28	
25	<211> 20	
	<212> ADN	
	<213> cebador	
30	<400> 28	
	cgttattayt ctayyycttc	20
35	<210> 29	
	<211> 20	
	<212> ADN	
40	<213> cebador	
	<400> 29	
45	gcmacaccrg tacctggacc	20
	<210> 30	
	<211> 18	
50	<212> ADN	
	<213> cebador	
	<400> 30	
55	atccaatcg taatcagc	18
	<210> 31	
60	<211> 18	
	<212> ADN	
	<213> cebador	
65	<400> 31	
	acttgtcttc gtttagca	18

ES 2 298 276 T3

	<210> 32	
	<211> 18	
	<212> ADN	
5	<213> cebador	
	<400> 32	
10	ctacgtctgt ggtgatgc	18
	<210> 33	
	<211> 25	
15	<212> ADN	
	<213> cebador	
	<400> 33	
20	ctcgggaagc gcgccattgt gttgg	25
	<210> 34	
25	<211> 29	
	<212> ADN	
	<213> cebador	
30	<400> 34	
	taatacgact cactataggg cgaattggc	29
35	<210> 35	
	<211> 30	
	<212> ADN	
40	<213> cebador	
	<400> 35	
45	gggttaatta acatactca agcagtttgg	30
	<210> 36	
	<211> 30	
50	<212> ADN	
	<213> cebador	
	<400> 36	
55	cccttaatta aggggggatg gaagtggcgcg	30
	<210> 37	
60	<211> 39	
	<212> ADN	
	<213> cebador	
65	<400> 37	
	ataagaatgc ggccgctgaa cgagaaccac atggaggag	39

ES 2 298 276 T3

	<210> 38	
	<211> 33	
	<212> ADN	
5	<213> cebador	
	<400> 38	
10	ccttaattaa ggataaccac atccatacgt cgc	33
	<210> 39	
	<211> 23	
15	<212> ADN	
	<213> cebador	
	<400> 39	
20	cacaccaccc acgacgactt gtg	23
	<210> 40	
25	<211> 21	
	<212> ADN	
	<213> cebador	
30	<400> 40	
	cttccgtgct gaacgactgc g	21
35	<210> 41	
	<211> 46	
	<212> ADN	
40	<213> cebador	
	<400> 41	
45	taatacgact cactataggg aggcacacca cccacgacga cttgtg	46
	<210> 42	
	<211> 42	
	<212> ADN	
50	<213> cebador	
	<400> 42	
55	cttccgtgct gaacgactgc gaattcttagc gcccttcaag tt	42
	<210> 43	
60	<211> 222	
	<212> ADN	
	<213> vector	
65		

ES 2 298 276 T3

<400> 43

```

5      caggsaacag ctatgaccat gattacgcca agcttggtag cgagctcaga tccactagta      60
      acggccgcga gtgtgctgga attcgccctt aagggcgaat tctgcagata tccatcacac      120
      tggcggccgc tcgagcatgc atctagaggg cccaattcgc cctatagtga gtcgtattac      180
10     aattcactgg ccgtcgtttt caaacgtcgt gactgggaaa ac                        222

```

<210> 44

<211> 222

15 <212> ADN

<213> vector

<400> 44

```

20     gtcccttggc gatactggtc ctatgagggt tcgaaccatg gctcggagct aggtgatcat      60
      tgcggggcgt cacacgacct taagccggaa ttcccgctta agcgtctat aggtagtgtg      120
25     accgcggcgg agctcgtacg tagatctccc gggttaagcg ggatatcact cagcataatg      180
      ttaagtgacc ggcagcaaaa tgttgacgca ctgacccttt tg                        222

```

30 <210> 45

<211> 1712

<212> ADN

35 <213> *Candida tropicalis*

<400> 45

```

40     ggtaccgagc tcacgagttt tgggattttc gagtttggat tgtttccttt gttgattgaa      60
      ttgacgaac cagaggtttt caagacagat aagattgggt ttatcaaac gcagtttgaa      120
45     atattccagt tggtttccaa gatattctga agaagattga cgatttgaaa ttgaagaag      180

```

50

55

60

65

ES 2 298 276 T3

	tggagaagat ctggttttga ttgttggaga atttcaagaa tctcaagatt tactctaacg	240
	acgggtacaa cgagaattgt attgaattga tcaagaacat gatcttggto ttacagaaca	300
5	tcaagttcct ggaccagact gagaatgcc aagatataca aggcgtcatg tgataaaatg	360
	gatgagattt atcccacaat tgaagaaaga gtttatggaa agtggtaaac cagaagctaa	420
10	acaggaagaa gcaaacgaag aggtcaaca agaagaagaa ggtaataaag tattttgtat	480
	tatata'acaa ncaaagtaag gaatacagat ttatacaata aattgcaata ctagtccagt	540
	gagatatctc atccattccc caactcccaa gaasaaaaaa aagtgaaaaa asaaatcaaa	600
15	cccaaagatc aacctcccca tcatcatcgt catcaaaccc ccagctcaat tcgcaatggt	660
	tagcacaaaa acatacacag acagggcctc agcacacccc tecaagggtg cccaaagttt	720
	attccgctta atggagttca aaaaagaccas cctctgcgac tcgacgcagc tgaccacaac	780
20	cgcgaggttc ctctcgctca tcgacaagct cggctcccaac atctgtctcg tgaagacgca	840
	catcgatata atctcagact tcagotacga gggcagcatt gagccgttgc ttgtccttgc	900
25	agagcgcac gggttcttga tattcgagga caggaagttt gctgatatcg gaacaccgt	960
	gatgttgcag tacacctcgg gggatatacg gatcgccggg tggagtgaac tcaagaacgc	1020
	gcacgagtg actgqcaagc gcctcgttga agggttgaaa cgcggtgcgg agggggtaga	1080
30	aaaggaaagg ggcgttctga tcttgccgga gttgtcgagt aaaggtcgtg tggcgcatgg	1140
	tgaatatacc cgtgagacga tcgagattgc gaagagtgat cgggagttcg tgattgggtt	1200
	catcgccgag cgggacatgg ggggtagaga agaaggggtt gattggatca tcatgacgcc	1260
35	tgggtgctggg ttggatgata aaggcgatgc gttgggccag cagtatagga ctgttgatga	1320
	ggtcgttctg actggtaccc atgtgattat tctcgggaga gggttgtttg gaaaaggaa	1380
40	agacctgag gtggaggga agagatacag ggtgctgga tggnaaggcat acttgaagag	1440
	aactggtcag ttgaataaaa tattgtaata aataggctcta tatacataca ctaagcttct	1500
	aggacgtcat tgtagtcttc gaagttgtct gctagtttag ttctcatgat ttcgaaaccc	1560
45	aataacgcaa tggatgtagc agggatggtg gttagtgcgt tccatgacaa cccagagtac	1620
	gccgcctcaa accacgtcac attcgcctt tgcctcatcc gcctcacttg ctggaaggta	1680
	tccacgtacg agttgtaata cccctgaag aa	1712
50		

LISTADO DE SECUENCIAS

- 55 <110> David, Craft L.
Madduri, Krishna
Loper, John C.
- 60 <120> GEN b5 DE CITOCROMO Y PROTEÍNAS DE *CANDIDA TROPICALIS* Y PROCEDIMIENTOS CORRESPONDIENTES
- <130> M6368 (1010-35 PCT)
- 65 <140> PCT/US 01/233376
- <141> 2001-07-25

ES 2 298 276 T3

<160> 45

<170> PatentIn version 3.1

5

<210> 1

<211> 2710

<212> ADN

10 <213> *Candida tropicalis*

<400> 1

15	acatacttca aqcaqtttgg cgacatagtg aacctcaagt tatcacaggaa caagacgacg	60
	ggcaagagca agcactacgg gtttatagag ttcacgtcgc ctgaagttag ccagatcgcg	120
	gcgagagcga tgaacaacta cttgtttggt ggacacttga tcaaaigtga ggttgcacgc	180
20	gagcgttca aggaactggt caaggactcg aagaggaagt tcaagtgat tccctggaag	240
	aagatcgcga aggataagca cgataagcca aagtcgcga aggagtggc gaagttggtg	300
	gagaagttag aagatccaa gaagaagaag caggaggagt tgaagagtaa aggtattgat	360
25	tttgatttgg ctgctatata aaggagataa gaggaggaga tgcacagcgc aaacgagcat	420
	tctgtgatg tgaagcag gtatagataa tagcggataa cgtaaaataa gagatctcca	480
	acttcaact tcaaaactcc gacccatc ttttggggga gagggttgg tatgtagtgg	540
30	tgaaggagag gaggatatt tgttttgcct aattgggata aattatccca gtcagttgaa	600
	agagcagagc gtaagccatt tctttttcta actgcaaata gcatacagat gcgatatgta	660
35	acgaagagag aatcaagag caggtgacta caticataga tagtgacatt ataataacat	720
	ggcgcatcat tgyttctatg tagctggcag ggttattatc aagcttgaat agtttaataa	780
	aatcgtacc atgaatgat gcatagaagc aataaggag cctgtgcctg tgagttagtag	840
40	cagtacggg gggagacgct agtttagggg taaaatgtca gcacatgaac agcagttgaa	900
	gtgggtgcca atcaagtaag aacatcttgt gaaaaatcaa aagcaatggt atatgtgttc	960
45	ctgcatacag tgctggagtc aacgagccaa aaaaaaaaaa gaaagaaaga gagaaaaact	1020

50

55

60

65

ES 2 298 276 T3

tatcgtataa aaaccacaca aaaatttccc aatcccaatt ccttcattct tcttctttta 1080
 ctgatttaac ccacagatac atacaattat gaccgacaca gacaccacga ccaccatcta 1140
 5 caccacagaa gaggttgccc agcacaccac ccacgacgac ttgtgggtta ttctcaatgg 1200
 taaggctctac aacatctcca actatataga cgagcaccga ggtggtgaag aagtcattct 1260
 10 tgattgcgcc ggcacagacg ccactgaagc ctttgacgac attggccact ccgacgaggc 1320
 ccacgagatc ttgaaaaagt tgtacattgg taacttgaag ggcgctaaga ttgttgaggc 1380
 15 caagcacgag cagtcgttca gcacggaaga agactcgggt atcaacttc cattgatctg 1440
 tggtgggtctg tttttggctg ctttcgggtg ctactactac aagaccaact ttgcctaage 1500
 ataacaagca gtacagttga aggacagggt agaggagatg agaaaaaacg ggaacccaac 1560
 20 aaagattatt ttcacacatc acatggaggg gctgatccca ctttttgacg tcaatatcca 1620
 cagcacgaag aaagaaagaa agaaagaaag tctatggaag aggaaatgga tcacattaga 1680
 25 gcttttcttt atgtaacata tatatatata taaactaata cagatttaca gatcacccac 1740
 atcaccgcag ggcttatcat ctgatggtgc ccaaaaaaaaa aaatccactg tggatgagcc 1800
 tagttaggag atatcgaggt agtcattct tttgatatct aggtcttcct ctcttgatt 1860
 ctacgttctg acttggtgct acacgatgag atcaccaggt gtcattctgg agtttggtgg 1920
 30 aaagtgtctt gattttttta gtaagcaaga atttggtgag ttctattgga tgttctggtg 1980
 35 cggccacttc catccccca ccccttgctt tgtcttgctt tgtcttattt ttttgggtcg 2040
 gttggcgcaa gtaagacgca cgcacaggag gagcacgacg gataaatatc cacttttttc 2100
 40 acacgcgtcg attgacggct tgtgtgaatt gtggggaata cggataaggg ggtataccac 2160
 acacacacat atctaacata tcagaccact ttctataaca gatctcatga tccccttgag 2220
 45 agttgatcca agtctatgt cctgtgatat tgccccccc ccccaagga agggcggggc 2280
 atgttatcag ggacctggat gaacccttga tggcggtgtg agtagatgca agagaggttg 2340
 tgctttgcaa gtagctgaag gtgtagggac atccggtact atagttctct tgaaggatca 2400
 50 tgccagctcc ctttctgtgg ctctctggaa gctctgcac ttctctctgt tgaacagcg 2460
 tggagttacg aaaggtaacc tgtggtgagt tcaacaaga catggctcta caagctgtcg 2520
 55 aggataazag taattaaaca acatgtatat atattaataa acggatccgt ggtgctagat 2580
 tgtggtagat gtttagtacc gtttatccc tctagtgaag actagcattt gattccatka 2640
 60 gtcacacta cttgatgta cattcaacca aatgaaggte ggtccaagat ccaagaatt 2700
 caaaaagctt 2710

<210> 2
 <211> 129

ES 2 298 276 T3

<212> PRT

<213> *Candida tropicalis*

5 <400> 2

```

Met Thr Asp Thr Asp Thr Thr Thr Thr Ile Tyr Thr His Glu Glu Val
1          5          10          15

Ala Gln His Thr Thr His Asp Asp Leu Trp Val Ile Leu Asn Gly Lys
20          25          30

Val Tyr Asn Ile Ser Asn Tyr Ile Asp Glu His Pro Gly Gly Glu Glu
35          40          45

Val Ile Leu Asp Cys Ala Gly Thr Asp Ala Thr Glu Ala Phe Asp Asp
50          55          60

Ile Gly His Ser Asp Glu Ala His Glu Ile Leu Glu Lys Leu Tyr Ile
65          70          75          80

Gly Asn Leu Lys Gly Ala Lys Ile Val Glu Ala Lys His Ala Gln Ser
85          90          95

Phe Ser Thr Glu Glu Asp Ser Gly Ile Asn Phe Pro Leu Ile Ala Val
100         105         110

Gly Val Phe Leu Ala Ala Phe Gly Val Tyr Tyr Tyr Lys Thr Asn Phe
115         120         125

Ala

```

<210> 3

<211> 4206

45 <212> ADN

<213> *Candida tropicalis*

<400> 3

```

50 catcaagatc atctatgggg ataattacga cagcaacatt gcagaaagag cggttggtcac 60

aatcgaaaga gcctatggcg ttgccgtcgt tgaggcaaat gacagcacca acaataacga 120

55 tgggtcccagt gaagagcctt cagaacagtc cattgttgac gcttaaggca cggataatta 180

```

60

65

ES 2 298 276 T3

cgtggggcaa aggaacgcgg aattagttat ggggggatca aaagcggaag atttgtgttg 240
 ettgtgggtt ttttccttta tttttcatat gatttctttg cgcaagtaac atgtgccaat 300
 5 ttagtttgtg attagcgtgc cccacaattg gcatcgtgga cgggcgtgtt ttgtcatacc 360
 ccaagtctta actagctcca cagtctcgac ggtgtctcga cgatgtcttc ttccacccct 420
 10 cccatgaatc attcaaagtt gttgggggat ctccaccaag ggcaccggag ttaatgctta 480
 tgtttctccc actttgggtg tgattggggg agtctagtga gttggagatt ttcttttttt 540
 15 cgcagggtgtc tccgatatcg aaatttgatg aatatagaga gaagccagat cagcacagta 600
 gattgccttt gtagttagag atgttgaaca gcaactagtt gaattacacg ccaccacttg 660
 acagcaagtg cagtgagctg taaacgatgc agccagagtg tcaccaccaa ctgacgttgg 720
 20 gtggagttgt tgttgttgtt gttggcaggg ccatattgct aaacgaagac aagtagcaca 780
 aaaccaagc ttaagaacaa aaataaaaaa aattcatcag acaattccaa agccattgat 840
 25 ttacataatc aacagtaaga cagaaaaaac ttcaacatt tcaaagttcc ctttttctta 900
 ttacttcttt ttttcttct ttcttcttt ccttctgtt ttcttacttt atcagtcttt 960
 30 tacttgtttt tgcaattcct catcctctc ctactctcc tcaccatggc tttagacaag 1020
 ttagatttgt atgtcatcat aacattgggt gtgcctgtag ccgcctattt tgetaagaac 1080
 cagttccttg atcagcccca ggacaccggg ttcttcaaca cggacagcgg aagcaactcc 1140
 35 agagacgtct tgctgacatt gaagaagaat aataaaaaca cgttggtgtt gtttgggtcc 1200
 cagacgggta cggcagaaga ttacgccaac aaattgtcca gagaattgca ctccagattt 1260
 40 ggcttgaaaa cgatggttgc agatttcgct gattacgatt gggataactt cggagatatc 1320
 accgaagaca tcttgggtgt tttcattgtt gccacctatg gtgagggtga acctaccgat 1380
 45 aatgcccagc agttccacac ctggttgact gaagaagctg acactttgag taccttgaaa 1440
 tacaccgtgt tcgggttggg taactccacg tacgagttct tcaatgccat tggtagaaa 1500
 50 tttgacagat tgttgagcga gaaagggtgt gacaggtttg ctgaatacgc tgaagggtgat 1560
 gacggtactg gcaccttgga cgaagatttc atggcctgga aggacaatgt ctttgacgcc 1620
 ttgaagaatg atttgaactt tgaagaaaag gaattgaagt acgaaccaa cgtgaaattg 1680
 55 actgagagag acgacttgct tgctgctgac tcccaagttt ccttgggtga gccaaacaag 1740
 aagtacatca actccgaggg catcgacttg accaagggtc cattcgacca caccaccca 1800
 60 tacttggcca gaatcaccca gacgagagag ttgttcagct ccaaggacag acactgtatc 1860

65

ES 2 298 276 T3

cacgttgaat ttgacatttc tgaatcgaac ttgaaataca ccaccggtga ccacttagct 1920
 atctggccat ccaactccga cgaaaacatt aagcaatttg ccaagtgttt cggattggaa 1980
 5 gataaactcg acactgttat tgaattgaag gcgttggact ccacttacac catcccattc 2040
 ccaaccccaa ttacctacgg tgctgtcatt agacaccatt tagaaatctc cggtcacgac 2100
 10 tcgagacaat tctttttgtc aatttgtggtg ttgtctctg atgaagaaac aaagaaggct 2160
 ttaccagac ttggtggtga caagcaagaa ttgcgcgcca aggtcacccg cagaaagttc 2220
 15 aacattgccg atgccttgtt atattctctc aacaacgctc catggtccga tgttcccttt 2280
 gaattcctta ttgaaaacgt tccacacttg actccacgtt actactccat ttcgtcttcg 2340
 tcattgagtg aaaagcaact catcaacgtt actgcagttg ttgaagccga agaagaagct 2400
 20 gatggcagac cagtcactgg tgttgtcacc aacttgttga agaacgttga aattgtgcaa 2460
 aacaagactg gcgaaaagcc acttgtccac tacgatttga gcggcccaag aggcaagttc 2520
 25 aacaagttca agttgccagt gcattgtgaga agatccaaact ttaagttgcc aaagaactcc 2580
 accacccacg ttatcttgat tggtcacggt actggtgttg cccattgag aggttttgtc 2640
 30 agagaaagag ttcaacaagt caagaatggt gtcaatgttg gcaagacttt gttgttttat 2700
 ggttgacaga actccaacga ggactttttg tacaagcaag aatgggccga gtacgcttct 2760
 35 gttttgggtg aaaactttga gatgttcaat gccttctcca gacaagacc atccaagaag 2820
 gtttacgtcc aggataagat tttagaaaac agccaacttg tgcacgagtt gttgactgaa 2880
 ggtgccatta tctacgtctg tggtagtgc agtagaatgg ctagagacgt gcagaccaca 2940
 40 atttccaaga ttgttgctaa aagcagagaa attagtgaag acaaggctgc tgaattggtc 3000
 aagtccttga aggtccaaaa tagataccaa gaagatgttt ggtagactca aacgaatctc 3060
 45 tcttctctcc aacgcattta tgaatcttta ttctcattga agctttacat atgttctaca 3120
 ctttattttt tttttttttt ttattattat attacgaac ataggtaac tatatatact 3180
 50 tgattaaatg ttatagaaac aataactatt atctactcgt ctacttcttt ggcattgaca 3240
 tcaacattac cgttccatt accgttgccg ttggcaatgc cgggatattt agtacagtat 3300
 55 ctccaatccg gatttgagct attgtagatc agctgcaagt cattctccac cttcaaccag 3360
 tacttatact tcatctttga cttcaagtcc aagtcataaa tattacaagt tagcaagaac 3420
 ttctggccat ccacgatata gacgttattc acgttattat gcgacgtatg gatgtggtta 3480
 60 tctttattga acttctcaaa cttcaaaaac aacccacgt cccgcaacgt cattatcaac 3540
 gacaagttct ggtcacgtc gtcggagctc gtcaagttct caattagatc gttcttqta 3600
 65

ES 2 298 276 T3

ttgatcttct ggtactttct caattgctgg aacacattgt cctcgtttgt caaatagatc 3660
 ttgaacaact ttttcaacgg gatcaacttc tcaatctggg ccaagatctc cgcctgggac 3720
 5 ttcagaaaca agtctctgca cccctggctg atggtctccg ggtacaacaa gtccaagggg 3780
 cagaagtgtc taggcacgtg tttcaactgg ttcaacgaac atgttcgaca gtagttcgag 3840
 10 ttatagttat cgtacaacca ttttggtttg atttcgaaaa tgacggagct gatgccatca 3900
 ttctctctgt tccctctata gtacaactgg caattcttcg agaggtctca ttcctctga 3960
 15 ttcccgctca agatattcgg caacaagagc cgtaccgtc cagggagcat caagtcgtgg 4020
 cccctggtgt tcaacttgt gatgaagtcc gaggtcaaga caatcaactg gatgtcgatg 4080
 atctggtgcg ggaacaagtt ctgcatctt agctcgatga agtcgtacaa ctccacacgc 4140
 20 gagatatact cctgttctc ctccaagagc cggatccgca agagcttgtg ctccaagtag 4200
 tcgttg 4206

25 <210> 4
 <211> 679
 <212> PRT
 30 <213> *Candida tropicalis*
 <400> 4

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

65

ES 2 298 276 T3

Phe Ile Val Ala Thr Tyr Gly Glu Gly Glu Pro Thr Asp Asn Ala Asp
115 120 125

Glu Phe His Thr Trp Leu Thr Glu Glu Ala Asp Thr Leu Ser Thr Leu
130 135 140

Lys Tyr Thr Val Phe Gly Leu Gly Asn Ser Thr Tyr Glu Phe Phe Asn
145 150 155 160

Ala Ile Gly Arg Lys Phe Asp Arg Leu Leu Ser Glu Lys Gly Gly Asp
165 170 175

Arg Phe Ala Glu Tyr Ala Glu Gly Asp Asp Gly Thr Gly Thr Leu Asp
180 185 190

Glu Asp Phe Met Ala Trp Lys Asp Asn Val Phe Asp Ala Leu Lys Asn
195 200 205

Asp Leu Asn Phe Glu Glu Lys Glu Leu Lys Tyr Glu Pro Asn Val Lys
210 215 220

Leu Thr Glu Arg Asp Asp Leu Ser Ala Ala Asp Ser Gln Val Ser Leu
225 230 235 240

Gly Glu Pro Asn Lys Lys Tyr Ile Asn Ser Glu Gly Ile Asp Leu Thr
245 250 255

Lys Gly Pro Phe Asp His Thr His Pro Tyr Leu Ala Arg Ile Thr Glu
260 265 270

Thr Arg Glu Leu Phe Ser Ser Lys Asp Arg His Cys Ile His Val Glu
275 280 285

Phe Asp Ile Ser Glu Ser Asn Leu Lys Tyr Thr Thr Gly Asp His Leu
290 295 300

Ala Ile Trp Pro Ser Asn Ser Asp Glu Asn Ile Lys Gln Phe Ala Lys
305 310 315 320

Cys Phe Gly Leu Glu Asp Lys Leu Asp Thr Val Ile Glu Leu Lys Ala
325 330 335

ES 2 298 276 T3

Leu Asp Ser Thr Tyr Thr Ile Pro Phe Pro Thr Pro Ile Thr Tyr Gly
340 345 350

Ala Val Ile Arg His His Leu Glu Ile Ser Gly Pro Val Ser Arg Gln
355 360 365

Phe Phe Leu Ser Ile Ala Gly Phe Ala Pro Asp Glu Glu Thr Lys Lys
370 375 380

Ala Phe Thr Arg Leu Gly Gly Asp Lys Gln Glu Phe Ala Ala Lys Val
385 390 395 400

Thr Arg Arg Lys Phe Asn Ile Ala Asp Ala Leu Leu Tyr Ser Ser Asn
405 410 415

Asn Ala Pro Trp Ser Asp Val Pro Phe Glu Phe Leu Ile Glu Asn Val
420 425 430

Pro His Leu Thr Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Ser Leu Ser
435 440 445

Glu Lys Gln Leu Ile Asn Val Thr Ala Val Val Glu Ala Glu Glu Glu
450 455 460

Ala Asp Gly Arg Pro Val Thr Gly Val Val Thr Asn Leu Leu Lys Asn
465 470 475 480

Val Glu Ile Val Gln Asn Lys Thr Gly Glu Lys Pro Leu Val His Tyr
485 490 495

Asp Leu Ser Gly Pro Arg Gly Lys Phe Asn Lys Phe Lys Leu Pro Val
500 505 510

His Val Arg Arg Ser Asn Phe Lys Leu Pro Lys Asn Ser Thr Thr Pro
515 520 525

Val Ile Leu Ile Gly Pro Gly Thr Gly Val Ala Pro Leu Arg Gly Phe
530 535 540

Val Arg Glu Arg Val Gln Gln Val Lys Asn Gly Val Asn Val Gly Lys
545 550 555 560

ES 2 298 276 T3

Thr Leu Leu Phe Tyr Gly Cys Arg Asn Ser Asn Glu Asp Phe Leu Tyr
565 570 575

Lys Gln Glu Trp Ala Glu Tyr Ala Ser Val Leu Gly Glu Asn Phe Glu
580 585 590

Met Phe Asn Ala Phe Ser Arg Gln Asp Pro Ser Lys Lys Val Tyr Val
595 600 605

Gln Asp Lys Ile Leu Glu Asn Ser Gln Leu Val His Glu Leu Leu Thr
610 615 620

Glu Gly Ala Ile Ile Tyr Val Cys Gly Asp Ala Ser Arg Met Ala Arg
625 630 635 640

Asp Val Gln Thr Thr Ile Ser Lys Ile Val Ala Lys Ser Arg Glu Ile
645 650 655

Ser Glu Asp Lys Ala Ala Glu Leu Val Lys Ser Trp Lys Val Gln Asn
660 665 670

Arg Tyr Gln Glu Asp Val Trp
675

<210> 5

<211> 4145

<212> ADN

<213> *Candida tropicalis*

<400> 5

tatatgatat atgatatatc ttcctgtgta attattattc gtattcgtta atacttacta	60
catttttttt tctttattta tgaagaaaag gagagttcgt aagttgagtt gagtagaata	120
ggctgttgtg catacgggga gcagaggaga gtatccgacg aggaggaact gggtgaaatt	180
tcatctatgc tgttgcgcc tgtactgtac tgtaaactct agatttccta gaggtgttc	240
tagcaataa agtggttcaa gatacaattt tacaggcaag ggtaaaggat caactgatta	300
gcggaagatt ggtgttgcc gtgggggtct tttatttttc atatgatttc ttgacgag	360
taacatgtgc caatctagtt tatgattagc gtacctccac aattggcatc ttggacgggc	420
gtgttttgc ttaccccaag ctttatttag ttccacagtc tcgacggtgt ctgcgcgatg	480
tcttctccca cccctgcgag gaatcattcg aagttgttg gggatctct ccgcagttta	540
tgttcatgtc tttccactt tgggtgtgat tggggtagcg tagtgagtg gtgattttct	600

ES 2 298 276 T3

	tttttcgcag gtgtctccga tategaagtt tgatgaatat aggagccaga tcagcatggt	660
	atattgcott tgtagataga gatgttgaac aacaactagc tgaattacac accaccgcta	720
5	aacgatgogc acaggggtgc accgccaaact gacgttgggt ggagttgttg ttggcagggc	780
	catattgcta aacgaagaga agtagcacia aaccaaggt taagaacaat taaaaaaatt	840
10	catacgacaa ttccacagcc atttacataa tcaacagcga caaatgagac agaaaaaact	900
	ttcaacattt caaagttccc tttttcttat tactttcttt ttttttctt tcttttcatt	960
15	tcttttcttt ctgcttttat tactttacca gtcttttgct tgtttttgca attcctcacc	1020
	ctcctcctca ccatggcttt agacaagtta gattttgtatg tcatcataac attgggtggc	1080
20	gctgtggccg cctattttgc taagaaccag ttctttgatc agccccagga caccgggttc	1140
	ctcaacacgg acagcggag caactccaga gacgtcttgc tgacattgaa gaagaataat	1200
	aaaaacacgt tgttgttgtt tgggtcccag accggtacgg cagaagatta cgccaacaaa	1260
25	ttgtcaagag aattgcactc cagatttggc ttgaaaacca tggttgcaga tttcgctgat	1320
	tacgattggg ataacttcgg agatatcacc gaagatatct tgggtgtttt catcggtgcc	1380
30	acctacgggt aggggtgaacc taccgacaat gccgacgagt tccacacctg gttgactgaa	1440
	gaagctgaca ctttgagtac tttgagatat accgtgttcg ggttgggtaa ctccacctac	1500
35	gagttcttca atgctatttg tagaaagttt gacagattgt tgagtgagaa aggtgggtgac	1560
	agatttgctg aatatgctga aggtgacgac ggcaactggc ccttggacga agatttcattg	1620
	gcttgggaagg ataattgtct tgacgccttg aagaatgact tgaactttga agaaaaggaa	1680
40	ttgaagtacg aaccaaacgt gaaattgact gagagagatg acttgtctgc tgccgactcc	1740
	caagtttctt tgggtgagcc aaacaagaag tacatcaact ccgagggcat cgacttgacc	1800
45	aagggtccat tcgaccacac ccaccatac ttggccagga tcaccgagac cagagagttg	1860
	ttcagctcca aggaaagaca ctgtattcac gttgaatttg acattttctga atcgaacttg	1920
50	aaatacacca ccggtgacca tctagccatc tggccatcca actccgacga aaacatcaag	1980
	caatttgcca agtgtttcgg attggaagat aaactcgaca ctgttattga attgaaggca	2040
	ttggactcca cttacaccat tccattccca actccaatta cttacgggtgc tgtcattaga	2100
55	caccatttag aaatctccgg tccagtctcg agacaattct tttgtcgat tgctgggttt	2160
	gtcctgatg aagaaacaaa gaagactttc accagacttg gtggtgacaa acaagaattc	2220
60	gccaccaagg ttaccgcag aaagttcaac attgccgatg ccttgttata ttcctccaac	2280

ES 2 298 276 T3

aacactccat ggtccgatgt tccttttgag ttccttattg aaaacatcca acacttgact 2340
 ccacgttact actccatttc tcttcgtcgt ttgagtgaag aacaactcat caatgttact 2400
 5 gcagtcgttg aggccgaaga agaagccgat ggcagaccag tcactggtgt tgttaccaac 2460
 ttgttgaaga acattgaaat tgcgcaaaac aagactggcg aaaagccact tgttcactac 2520
 10 gatttgagcg gcccaagagg caagttcaac aagttcaagt tgccagtga cgtgagaaga 2580
 tccaacttta agttgccaaa gaactccacc accccagtta tcttgattgg tccaggtact 2640
 15 ggtgttgccc cattgagagg tctcgttaga gaaagagttc aacaagtcaa gaatggtgtc 2700
 aatgttgcca agactttgtt gttttatggt tgcagaaact ccaacgagga ctttttgtac 2760
 aagcaagaat gggccgagta cgcttctgtt ttgggtgaag actttgagat gttcaatgcc 2820
 20 tctctagac aagaccctac caagaagggt taogtccagg ataagatttt agaaaacagc 2880
 caacttgtgc acgaattgtt gaccgaaggt gccattatct acgtctgtgg tgacgccagt 2940
 25 agaattggcca gagacgtcca gaccacgac tcacaagattg ttgccaaaag cagagaaatc 3000
 agtgaagaca aggccgctga attggtcaag tcctggaaag tccaaaatag ataccaagaa 3060
 30 gatgtttggt agactcaaac gaatctctct tctcccaac gcatttatga atattctcat 3120
 tgaagtttta catatgttct atatttcatt tttttttat tatattacga aacatagggtc 3180
 aactatatat acttgattaa atgttataga aacaataatt attatctact cgtctacttc 3240
 35 tttggcattg gcattggcat tggcattggc attgcccgtg ccgttggtta tgcgggata 3300
 tttagtacag tatctccaat ccggatttga gctattgtta atcagctga agtcattctc 3360
 40 cacttcaac cagtacttat acttcattct tgacttcaag tccaagtcac aaatattaca 3420
 agttagcaag aacttctggc catccacaat atagacgtta ttcacgttat tatgcgacgt 3480
 45 atgatatgg ttatccttat tgaacttctc aaacttcaaa aacaacccca cgtcccgcga 3540
 cgtcattatc aacgacaagt tctgactcac gtggtcggag ctggtcaagt tctcaattag 3600
 atcgttcttg ttattgatct tctggtactt tctcaactgc tggaaacacat tgtcctcgtt 3660
 50 gttcaaatag atcttgaaca acttcttcaa gggaatcaac ttttcgatct gggccaagat 3720
 ttcgcccggg atcttcagaa acaagtcttg caacccttg tcgatggtct cggggtacaa 3780
 55 caagtctaag gggcagaagt gtctaggcac gtgtttcaac tggttcaagg aacatgttcg 3840
 acagtagttc gagttatagt tatcgtacaa ccactttggc ttgatttoga aaatgacgga 3900
 60 gctgatccca tcattctcct ggttccttcc atagtacaac tggcatttct tcgagagact 3960
 caactcctcg tagttcccg ccaagatatt cggcaacaag agcccgtagc gctcacggag 4020

ES 2 298 276 T3

```

catcaagtcg tggeccctggg tgttcaactt gttgatgaag tccgatgtca agacaatcaa 40
ctggatgtcg atgatctggg gccgaaacaa gttcttgcac tttagctcga tgaagtcgta 41
5      caact 41

```

<210> 6

<211> 679

<212> PRT

<213> *Candida tropicalis*

<220>

<221> MISC_FEATURE

<222> (50)..(50)

<223> Xaa = leucina o serina

<400> 6

```

Met Ala Leu Asp Lys Leu Asp Leu Tyr Val Ile Ile Thr Leu Val Val
1          5          10          15

```

```

Ala Val Ala Ala Tyr Phe Ala Lys Asn Gln Phe Leu Asp Gln Pro Gln
20          25          30

```

```

Asp Thr Gly Phe Leu Asn Thr Asp Ser Gly Ser Asn Ser Arg Asp Val
35          40          45

```

```

Leu Xaa Thr Leu Lys Lys Asn Asn Lys Asn Thr Leu Leu Leu Phe Gly
50          55          60

```

```

Ser Gln Thr Gly Thr Ala Glu Asp Tyr Ala Asn Lys Leu Ser Arg Glu
65          70          75          80

```

```

Leu His Ser Arg Phe Gly Leu Lys Thr Met Val Ala Asp Phe Ala Asp
85          90          95

```

```

Tyr Asp Trp Asp Asn Phe Gly Asp Ile Thr Glu Asp Ile Leu Val Phe
100          105          110

```

```

Phe Ile Val Ala Thr Tyr Gly Glu Gly Glu Pro Thr Asp Asn Ala Asp
115          120          125

```

```

Glu Phe His Thr Trp Leu Thr Glu Glu Ala Asp Thr Leu Ser Thr Leu
130          135          140

```

ES 2 298 276 T3

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

ES 2 298 276 T3

	370	375	380
5	Thr Phe Thr Arg Leu Gly Gly Asp Lys Gln Glu Phe Ala Thr Lys Val 385 390 395 400		
10	Thr Arg Arg Lys Phe Asn Ile Ala Asp Ala Leu Leu Tyr Ser Ser Asn 405 410 415		
15	Asn Thr Pro Trp Ser Asp Val Pro Phe Glu Phe Leu Ile Glu Asn Ile 420 425 430		
20	Gln His Leu Thr Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Ser Leu Ser 435 440 445		
25	Glu Lys Gln Leu Ile Asn Val Thr Ala Val Val Glu Ala Glu Glu Glu 450 455 460		
30	Ala Asp Gly Arg Pro Val Thr Gly Val Val Thr Asn Leu Leu Lys Asn 465 470 475 480		
35	Ile Glu Ile Ala Gln Asn Lys Thr Gly Glu Lys Pro Leu Val His Tyr 485 490 495		
40	Asp Leu Ser Gly Pro Arg Gly Lys Phe Asn Lys Phe Lys Leu Pro Val 500 505 510		
45	His Val Arg Arg Ser Asn Phe Lys Leu Pro Lys Asn Ser Thr Thr Pro 515 520 525		
50	Val Ile Leu Ile Gly Pro Gly Thr Gly Val Ala Pro Leu Arg Gly Phe 530 535 540		
55	Val Arg Glu Arg Val Gln Gln Val Lys Asn Gly Val Asn Val Gly Lys 545 550 555 560		
60	Thr Leu Leu Phe Tyr Gly Cys Arg Asn Ser Asn Glu Asp Phe Leu Tyr 565 570 575		
65	Lys Gln Glu Trp Ala Glu Tyr Ala Ser Val Leu Gly Glu Asn Phe Glu 580 585 590		
70	Met Phe Asn Ala Phe Ser Arg Gln Asp Pro Ser Lys Lys Val Tyr Val 595 600 605		

ES 2 298 276 T3

Gln Asp Lys Ile Leu Glu Asn Ser Gln Leu Val His Glu Leu Leu Thr
610 615 620

Glu Gly Ala Ile Ile Tyr Val Cys Gly Asp Ala Ser Arg Met Ala Arg
625 630 635 640

Asp Val Gln Thr Thr Ile Ser Lys Ile Val Ala Lys Ser Arg Glu Ile
645 650 655

Ser Glu Asp Lys Ala Ala Glu Leu Val Lys Ser Trp Lys Val Gln Asn
660 665 670

Arg Tyr Gln Glu Asp Val Trp
675

<210> 7

<211> 4115

<212> ADN

<213> *Candida tropicalis*

<400> 7

```

catatgcgct aatcttcttt tcttttttat cacaggagaa actatccac ccccacttcg      60
aacacaaatg acaactcctg cgttaacttg aaattcttgt ctgactaatt gaaaactccg      120
gacgagtcag acctccagtc aaacggacag acagacaaac acttggtgcg atgttcatac      180
ctacagacat gtcaacgggt gttagacgac ggtttcttgc aaagacaggt gttggcatct      240
cgtacgatgg caactgcagg aggtgtcgac ttctccttta ggcaatagaa aaagactaag      300
agaacagcgt ttttacaggt tgcattgggt aatgtagtat ttttttagtc ccagcattct      360
gtgggttgcg ctgggtttct agaataggaa atcacaggag cctgcaatt cagatggaag      420
aacaaagaga taaaaaacia aaaaaaactg agttttgcac caatagaatg tttgatgata      480
tcattccactc gctaaacgaa tcatgtgggt gatcttctct ttagttttgg tctatcataa      540
aacacatgaa agtgaatatc aaatacacta cactccgggt attgtccttc gttttacaga      600
tgtctcattg tcttactttt gaggtcatag gagttgcctg tgagagatca cagagattat      660
cacactcaca tttatcgtag ttctctatct catgctgtgt gtctctgggt gggttcagag      720
tttgatttgt tgtacattaa aggaatcgtt ggaaagcaaa gctaactaaa ttttctttgt      780
cacaggtaca ctaacctgta aaacttcact gccacgccag tctttcctga ttgggcaagt      840
gcacaaacta caacctgcaa aacagcactc cgcttgtcac aggttgtctc ctctcaacca      900

```

ES 2 298 276 T3

acaaaaaaat aagattaaac tttctttgct catgcatcaa teggagttat ctctgaaaga 960
 gttgcctttg tgtaatgtgt gccaaactca aactgcaaaa ctaaccacag aatgatttcc 1020
 5 ctcacaatta tataaactca cccacatttc cacagaccgt aatttcattgt ctccatttct 1080
 cttttgctct tcttttactt agtcagggtt gataacttcc ttttttatta cccatcttta 1140
 10 tttattttatt tattcattta taccaaccac ccaaccatgg ccacacaaga aatcatcgat 1200
 tctgtacttc cgtacttgac caaatggtac actgtgatta ctgcagcagt attagtcttc 1260
 15 cttatctoca caaacatcaa gaactacgtc aaggcaaga aattgaaatg tgtcgatcca 1320
 ccatacttga aggatgccgg tctcactggt attctgtctt tgatcgccgc catcaaggcc 1380
 aagaacgacg gtagattggc taactttgcc gatgaagttt tcgacgagta cccaaaccac 1440
 20 accttctact tgtctgttgc cggkgtcttg aagattgtca tgactgttga cccagaaaac 1500
 atcaaggctg tcttggccac ccaattcact gacttctcct tgggtaccag acacgcccac 1560
 25 tttgctcctt tgttgggtga cggkatcttc accttggacg gagaaggttg gaagcactcc 1620
 agagctatgt tgagaccaca gtttgctaga gaccagattg gacacgttaa agccttggaa 1680
 30 ccacacatcc aaatcatggc taagcagatc aagttgaacc agggaaagac tttcgatctc 1740
 caagaattgt tctttagatt taccgtcgac accgctactg agttcttgtt tgggtaatcc 1800
 gttcactcct tgtacgatga aaaattgggc atcccaactc caaacgaaat cccaggaaga 1860
 35 gaaaactttg ccgctgcttt caacggttcc caacactact tggccaccag aagttactcc 1920
 cagacttttt actttttgac caaccttaag gaattcagag actgtaacgc caaggtccac 1980
 40 cacttggcca agtactttgt caacaaggcc ttgaacttta ctctgaaga actcgaagag 2040
 aaatccaagt ccggttacgt tttcttgtac gaattggta agcaaccag agatccaaag 2100
 45 gtcttgcaag atcaattgtt gaacattatg gttgccggaa gagacaccac tgcgggtttg 2160
 ttgtcctttg ctttgtttga attggctaga caccagaga tgtgttcaa gttgagagaa 2220
 50 gaaatcgaag ttaacttttg tgttgggtga gactcccgcg ttgaagaaat taccttcgaa 2280
 gccttgaaga gatgtgaata cttgaaggct atccttaacg aaaccttgcg tatgtacca 2340
 tctgttcttg tcaactttag aaccgccacc agagacacca ctttgccaag aggtggtggc 2400
 55 gctaacggta ccgaccatct ctacattcct aaaggtcca ctgttgctta cgttgtctac 2460
 aagaccacc gtttgaaga atactacggt aaggacgcta acgacttcag accagaaaga 2520
 60 tggtttgaac catctactaa gaagttgggc tgggcttatg ttccattcaa cgggtgtcca 2580

ES 2 298 276 T3

agagtctgct tgggtcaaca attcgcccttg actgaagcct cttatgtgat cactagattg 2640
 gccagatgt ttgaaactgt ctcatctgat ccaggctctg aataccctcc accaaagtgt 2700
 5 attcacttga ccattgagtc caacgatggt gtctttgtca agatgtaaag tagtcgatgc 2760
 tgggtattcg attacatgtg tataggaaga ttttggtttt ctattcgttc ttttttttaa 2820
 10 tttttgttaa attagtttag agatttcatt aatacataga tgggtgctct ttccgaaact 2880
 ttactttctat cccctgtatc ccttattatc cctctcagtc acatgattgc tgtaattgtc 2940
 gtgcaggaca caaactccct aacggactta aaccataaac aagctcagaa ccataagccg 3000
 15 acatcactcc ttcttctctc ttstccaacc aatagcatgg acagaccac cctcttatcc 3060
 gaatcgaaga ccttattga ctccataccc acctggaagc cctcaagcc acacacgtca 3120
 20 tccagcccac ccataccac atccctctac tcgacaacgt ccaaagacgg cgagttctgg 3180
 tgtgcccgga aatcagccat ccggccaca tacaagcagc cgttgattgc gtgcatactc 3240
 ggcgagccca caatgggagc cagcattcg gaccatgaag caaagtacat tcacgagatc 3300
 25 acgggtgttt cagtgtcga gattgagaag ttcgacgatg gatggaagta cgatctcgtt 3360
 gccgattacg acttcgggtg gtgtttatct aaacgaagat tctatgagac gcagcatgtg 3420
 30 ttccggttcg aggattgtgc gtacgtcatg agtgtgcctt ttgatggacc caaggaggaa 3480
 ggttacgtgg ttgggacgta cagatccatt gaaaggttga gctygggtta agacggggac 3540
 35 gtggagtggg ccattggcag gacgtcggat cctgggtgggt ttatcccgca atggataact 3600
 cgattgagca tccctggagc aatcgcaaaa gatgtgccta gtgtattaaa ctacatacag 3660
 aaataaaaaac gtgtcttgat tcattgggtt ggttcttgtt gggttccgag ccaatatttc 3720
 40 acatcatctc ctcaattctc caagaatccc aacgtagcgt agtccagcac gccctctgag 3780
 atcttattta atatcgactt ctcaaccacc ggtggaatcc cgttcagacc attgttacct 3840
 45 gtagtgtgtt tgcctctgtt cttgatgaca atgatgtatt tgtcaagata cctgaaataa 3900
 taaaacatcc agtcattgag cttattactc gtgaacttat gaaagaactc attcaagccg 3960
 50 ttcccaaaaa acccagaatt gaagatcttg ctcaactggt catgcaagta gtagatcgcc 4020
 atgatctgat actttaccaa gctatcctct ccaagttctc ccacgtacgg caagtacggc 4080
 aacgagctct ggaagctttg ttgtttgggg tcata 4115

55

<210> 8

<211> 3948

60 <212> ADN

<213> *Candida tropicalis*

65

ES 2 298 276 T3

<400> 8

5	gacctgtgac gcttcgggtg tcttgccacc agtctccaag ttgaccgacg cccaagtcac	60
	gtaccacttt atttcgggtt acacttccaa gatggctggt actgaagaag gtgtcacgga	120
	accacaagct actttctcgg cttgtttcgg tcaaccattc ttgggtgttg acccaatgaa	180
10	gtacgctcaa caattgtctg acaagatctc gcaacacaag gctaacgcct gggtgttgaa	240
	caccgggttg gttggttctt ctgctgctag aggtggtaag agatgctcat tgaagtacac	300
15	cagagccatt ttggacgcta tccactctgg tgaattgtcc aaggttgaat acgaaacttt	360
	cccagtcttc aacttgaatg tcccaacctc ctgtccaggt gtcccaagtg aaatcttgaa	420
20	cccaaccaag gcctggaccg gaagggtgtt actccttcaa caaggaaatc aagtctttgg	480
	ctggttaagtt tgctgaaaac ttcaagacct atgctgacca agctaccgct gaagtgagag	540
	ctgcaggctc agaagcttaa agatatttat tcattattta gtttgccctat ttatttctca	600
25	ttacccatca tcattcaaca ctatatataa agtctcttcg gatatcattg taatcgtgcg	660
	tgtcgcaatt ggatgatttg gaactgcgct tgaacggat tcatgcacga agcggagata	720
30	aaagattacg taatttatct cctgagacaa ttttagccgt gttcacacgc ccttctttgt	780
	tctgagcgaa ggataaataa ttagacttcc acagctcatt ctaatttccg tcacgcgaat	840
35	attgaagggg ggtacatgtg gccgctgaat gtgggggcag taaacgcagt ctctctctc	900
	ccaggaatag tgcaacggag gaaggataac gtagagaaag cggaatgcga ggaaaatttt	960
	gaacgcgcaa gaaaagcaat atccgggcta ccaggttttg agccaggga cactctcta	1020
40	tttctgctca atgactgaac atagaaaaaa caccaagacg caatgaaacg cacatggaca	1080
	tttagacctc cccacatgtg atagtttgtc ttaacagaaa agtataataa gaacccatgc	1140
45	cgtccctttt ctttcgccgc ttcaactttt ttttttttat cttacacaca tcacgaccat	1200
	gactgtacac gatattatcg ccacatactt caccaaatgg tacgtgatag taccactcgc	1260
50	tttgattgct tatagagtcc tcgactactt ctatggcaga tacttgatgt acaagcttgg	1320
	tgctaaacca tttttccaga aacagacaga cggctgttcc ggattcaaag ctccgcttga	1380
55	attgttgaag aagaagagcg acggtaccct catagacttc acactccagc gtatccacga	1440
	tctcgatcgt cccgatatcc caactttcac attcccgtc ttttccatca accttgtcaa	1500
	taccttgag cgggagaaca tcaaggccat cttggccact cagttcaacg atttctcctt	1560
60	gggtaccaga cactcgcaact ttgctccttt gttgggtgat ggtatcttta cgttggatgg	1620
	cgccggctgg aagcacagca gatctatgtt gagaccacag ttgcccagag aacagatttc	1680

65

ES 2 298 276 T3

ccacgtcaag ttgttgagc cacacgttca ggtgttcttc aaacacgtca gaaaggcaca 1740

gggcaagact tttagacatcc aggaattgtt tttagagattg accgtcgact ccgccaccga 1800

5 gtttttgttt ggtgaatccg ttgagtcctt gagagatgaa tctatcggca tgtccatcaa 1860

tgcgcttgac tttagcggca aggcctggctt tgctgatgct tttactatt cgcagaatta 1920

10 tttagcttcg agagcgggta tgcaacaatt gtactgggtg ttgaacggga aaaagttaa 1980

ggagtgaac gctaaagtgc acaagtgtgc tgactactac gtcaacaagg ctttggaactt 2040

15 gacgcctgaa caattggaaa agcaggatgg ttatgtgttt ttgtacgaat tgggcaagca 2100

aaccagagac aagcaagtgt tgagagacca attgttgaac atcatggttg ctggtagaga 2160

caccaccgcc ggtttgttgt cgtttgtttt ctttgaattg gccagaaacc cagaagtac 2220

20 caacaagtig agagaagaaa ttgaggacaa gtttggactc ggtgagaatg ctagtgttga 2280

agacatttcc tttagatcgt tgaagtcctg tgaatacttg aaggctgttc tcaacgaac 2340

25 cttgagattg taccatccg tgccacagaa tttagagatt gccaccaaga acactaccct 2400

cccaagaggt ggtggttaagg acgggtgtgc tcctgttttg gtgagaaagg gtcagaccgt 2460

30 tatttacggt gtctacgag ccacagaaa ccagctgtt tacggttaagg acgtcttga 2520

gttagacca gagagatggt ttgagccaga gacaaagaag cttggtggg ccttctccc 2580

attcaacggt ggtcaagaa tctgtttggg acagcagtt gccttgacag aagcttcgta 2640

35 tgtcactgic aggttgctcc aggagttgc acattgtct atggaccag acaccgaata 2700

tccacctaag aaaatgtgc atttgaccat gtcgttttc gacggtgcc atattgagat 2760

40 gtattagagg gtcattgttt attttgattg tttagttgt aattactgat taggttaatt 2820

catggattgt tatttattga taggggttg cgcgtgttg attcacttg gatcgttcca 2880

ggttgatgtt tccttccatc ctgtcgagtc aaaaggagtt ttgttttgta actccggacg 2940

45 atgttttaa tagaaggctg atctccatgt gattgttttg actgttactg tgattatgta 3000

atctgcggac gttatacaag catgtgattg tggttttgca gccttttgca cgacaaatga 3060

50 tcgtcagacg attacgtaat cttgtttaga ggggtaaaaa aaaacaaat ggcagccaga 3120

atttcaaca ttctgcaaac aatgcaaaaa atgggaaact ccaacagaca aaaaaaaaaa 3180

55 ctccgcagca ctccgaaccc acagaacaat ggggcgcag aattattgac tattgtgact 3240

ttttacgt aacgtcatt gcagtgtagt gcgtcttaca cggggtattg ctttctaca 3300

60 tgcaagggca cagttgaagg tttagacctt acgttgcccc gtgtcaactc aatttgacga 3360

65

ES 2 298 276 T3

gtaacttect aagctcgaat tatgcagctc gtgcgtcaac ctatgtgcag gaaagaaaa 342
 atccaaaaaa atcgaaaatg cgacttttca ttttgaataa accaaaaaga aaaatgtcgc 348
 5 acttttttct cgctctcgct ctctcgaccc aaatcacaac aaatcctcgc gcgcagtatt 354
 tcgacgaaac cacaacaaat aaaaaaaca aattctacac caattctttt tcttcaccag 360
 10 tcaacaaaaa acaacaaatt atacaccatt tcaacgattt ttgctcttat aaatgctata 366
 taatggttta attcaactca ggtatgttta ttttactgtt ttcagctcaa gtatgttcaa 372
 atactaacta cttttgatgt ttgtcgcttt totagaatca aaacaacgcc cacaacacgc 378
 15 cgagcttgtc gaatagacgg tttgtttact cattagatgg tcccagatta cttttcaagc 384
 caaagtctct cgagttttgt ttgctgtttc cccaattcct aactatgaag ggtttttata 390
 20 aggtccaaag accccaaggc atagtctttt tggttccttc ttgtcgtg 394

<210> 9

25 <211> 3755

<212> ADN

<213> *Candida tropicalis*

30 <400> 9

gctcaacaat tgtctgacaa gatctcgcaa cacaaggcta acgcctgggt gttgaacact 60
 35 ggttggggtg gttcttctgc tgctagaggt ggtaagagat gttcattgaa gtacaccaga 120
 gccattttgg acgctatcca ctctggtgaa ttgtccaagg ttgaatacga gactttccca 180
 40 gtcttcaact tgaatgtccc aacctcctgc ccagggtgct caagtgaat cttgaaccca 240
 accaaggcct ggaccgaagg tgttgactcc ttcaacaagg aaatcaagtc tttggctggt 300
 aagtttgctg aaaacttcaa gacctatgct gaccaagcta cgcctgaagt tagagctgca 360
 45 ggtccagaag cttaaagata ttatttcact atttagtttg cctatttatt tctcatcacc 420
 catcatcatt caacaatata tataaagtta tttcggaact catatatcat tgtaatcgtg 480
 50 cgtgttgcaa ttgggtaatt tgsaactgta gttggaacgg attcatgcac gatgcggaga 540
 taacacgaga ttatctcta agtcaatttt ggctcattc acacgccctt ctctgagct 600
 aaggataaat aattagactt cacaagttca ttaaaatata cgtcacgga aaactgcaac 660
 55 aataagggaag yggggggtag acgtagccga tgaatgtggg gtgccagtaa acgcagcttc 720
 tctctcccc ccccccccc cccctcagg aatagtacaa cgggggaagg ataacggata 780
 60 gcaagtggaa tgcgaggaaa attttgaatg cgcaaggaaa gcaatatccg ggctatcagg 840
 ttttgagcca ggggacacac tctcttctg cacaanaact taacgtagac aaaaaaaaa 900

65

ES 2 298 276 T3

aactccacca agacacaatg aatcgacat ggacatttag acctcccccac atgtgaaagc 960
 ttctctggcg aaagcaaaaa aagtataata aggacccatg ccttccctct tcttgggccc 1020
 5 tttcaacttt ttctttttct ttgtctatca acacacacac acctcacgac catgactgca 1080
 caggatatta tgcacacata catcaccaaa tggtagctga tagtaccact cgctttgatt 1140
 10 gcttataggg tctctgacta cttttacggc agatacttga tgtacaagct tgggtgctaaa 1200
 ccgttttttc agaacaacac agacgggttat ttcggtatca aagctccact tgaattgtta 1260
 15 aaaaagaaga gtgacggtag cctcatagac ttactctctg agcgtatcca agcgtcaat 1320
 cgtccagata tcccaacttt tacattccca atcttttcca tcaaccttat cagcacctt 1380
 gagccggaga acatcaaggc tatcttggcc acccagttca acgatttctc cttgggcacc 1440
 20 agacactcgc actttgctcc ttgttgggc gatggtatct ttaccttga cgggtgcggc 1500
 tggaaqcaca gcagatctat gttgagacca cagtttgcca gagaacagat tccccacgtc 1560
 25 aagttgttgg agccacacat gcaggtgttc ttcaagcacg tcagaaaggc acagggcaag 1620
 acttttgaca tccaaagatt gtttttcaga ttgacgtcg actccgccac tgagtttttg 1680
 30 tttggtgaat ccgttgagtc cttgagagat gaactctattg ggatgtccat caatgcactt 1740
 gactttgacg gcaaggctgg ctttgcctgat gcttttaact actcgcagaa ctatttggtc 1800
 35 tctgagagcgg ttatgcaaca attgtactgg gtgttgaacg ggaaaaagt taaggagtgc 1860
 aacgctaaag tgcaaaagt ttgtgactat tacgtcagca aggcttttga cttgacacct 1920
 gaacaatttg aaaaacagga tggttatgtg ttcttgtacg agttggtcaa gcaaaccaga 1980
 40 gacaggcaag tgttgagaga ccagttgttg aacatcatgg ttgccggtag agacaccacc 2040
 gccggtttgt tgtctgtgt tttctttgaa ttggccagaa acccagaggt gaccaacaag 2100
 45 ttgagagaag aaatcgagga caagtttggc cttggtgaga atgctcgtgt tgaagacatt 2160
 tcttttgagt cgttgaagtc atgtgaatac ttgaaggctg ttctcaacga aactttgaga 2220
 50 ttgtacceat ccgtgacaca gaatttcaga gttgccacca aaaacactac ccttccaagg 2280
 ggaggttgta aggaagggtt atctcctgtt ttggtcagaa agggcacaac cgttatgtac 2340
 ggtgtctacg ctgcccacag aaaccagct gtctacggtt aggaagccct tgagtttga 2400
 55 ccagagaggt ggttgagcc agagacaaag aagcttggct gggccttctt tccattcaac 2460
 ggtggtccaa gaatttgctt gggacagcag ttgacctga cagaagcttc gtatgtcact 2520
 60 gtcagattgc tccagaggt ttgacaattg tctatggacc ccaacaccga atatccacct 2580
 aggaaaatgt cgcattgac catgtccctt ttcgacgggtg ccaacattga gatgtattag 2640
 65

ES 2 298 276 T3

5 aggatcatgt gttatTTTTg attggTTtag tctgTTtgta gctattgatt aggttaattc 2700
 acggattgtt atttattgat agggggTgcg tgtgtgtgtg tgtgttgcat tcacatggga 2760
 10 tcgTTccagg ttgTTgTTtc ctccatcct gttgagtcaa aaggagtttt gttttgtaac 2820
 tccggacgat gtcttagata gaaggTcgat ctccatgtga ttgTTtgact gctactctga 2880
 15 ttatgtaatc tgtaaagcct agacgttatg caagcatgtg attgtggTTt ttgcaacctg 2940
 tttgcacgac aaatgatcga cagtgcatta cgtaatccat attatttaga ggggtaataa 3000
 20 aaaataaatg gcagccagaa ttTcaaacat tttgcaaaca atgcaaaaga tgagaaactc 3060
 caacagaaaa aataaaaaaa ctccgcagca ctccgaacca acaaaacaat ggggggcgcc 3120
 agaattattg actattgtga ctttttttta tttttccgt taactttcat tgcagtgaag 3180
 25 tgtgttacac ggggtggtga tgggtgttgg ttctacaatg caagggcaca gttgaagggt 3240
 tcacataaac gttgcacct atcaactcaa tttatcctca ttcatgtgat aaaagaagag 3300
 30 ccaaaaggta attggcagac cccccaaggg gaacacggag tagaaagcaa tggaaacacg 3360
 cccatgacag tgccatttag cccacaacac atctagtatt cttttttttt tttgtgcga 3420
 ggtgcacacc tggacttttag ttattgcccc ataaagttaa caatctcacc tttggtctc 3480
 35 ccagtgtctc cgctccaga tgcctgtttt acacctcga gctaacgaca acacaacacc 3540
 catgagggga atgggcaaag ttaaacactt ttggtttcaa tgattcctat ttgctaactc 3600
 40 cttgttttgt gttttgattt gaacctgtg aaataaacga caattatata taacttttgc 3660
 tetgtctctc aatgtctctt ttgtgtgcca tttgtctttt tgctttttgc ttttgacctc 3720
 tctccactc ccacaatcag tgcagcaaca caca 3755

<210> 10

<211> 3900

<212> ADN

<213> *Candida tropicalis*

<400> 10

50 gacatcataa tgacccggtt atttcgcct caggttgctt atttgagccg taaagtgcag 60
 tagaaacttt gccttgggtt caaactctag tataatggtg ataactggtt gcactcttgc 120
 55 cataggcatg aaaataggcc gttatagtac tatatttaat aagcgtagga gtataggatg 180
 catatgaccg gttttctat atttttaaga taactctag taaattttgt attctcagta 240
 60 ggatttcac ccaatttcga accaattctg gcgaaaaaat gattctttta cgtcaaaagc 300
 tgaatagtgc agtttaaaac acctaaaac acatatacag cctctagata cgacagagaa 360

ES 2 298 276 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

gctctttatg atctgaagaa gcattagaat agctactatg agccactatt ggtgtatata 420
 ttagggattg gtgcaattaa gtacgtacta ataaacagaa gaaaatactt aaccaatttc 480
 tgggtgtatac ttagtgggtga gggacctttt ctgaacattc ggggtcaaact tttttttgga 540
 gtgcgacatc gatttttctgt ttgtgtaata atagtgaacc tttgtgtaat aaatcttcat 600
 gcaagacttg cataattcga gcttgggagt tcacgccaat ttgacctcgt tcatgtgata 660
 aaagaaaaagc caaaaggtaa ttagcagacg caatgggaac atggagtggg aagcaatgga 720
 agcacgcccc ggacggagta atttagtcca cactacatct ggggggtttt tttttgtgcg 780
 caagtacaca cctggacttt agtttttggc ccataaagtt aacaatctaa cctttggctc 840
 tccaactctc tccgccccca aatattcgtt ttacacctt caagctagcg acagcacaac 900
 acccattaga ggaatggggc aaagttaaac acttttggct tcaatgattc ctattcgcta 960
 ctacattctt ctcttgtttt gtgctttgaa ttgcaccatg tgaaataaac gacaattata 1020
 tatacctttt catccctcct cctatatctc tttttgttac attttgtttt ttacgtttct 1080
 tgcttttgca ctctcccact ccacaaaaga aaaaaaaact acactatgtc gtcttctcca 1140
 tcgtttgccc aagaggttct cgctaccact agtccttaca tcgagtactt tcttgacaac 1200
 tacaccagat ggtactactt catacctttg gtgcttcttt cgttgaactt tataagtttg 1260
 ctccacacaa ggtacttgga acgcaggttc cagccaagc cactcggtta cttgtcagg 1320
 gacctacgt ttggtatcgc tactccgttg cttttgatct acttgaagtc gaaaggtagc 1380
 gtcatgaagt ttgcttgggg cctctggaac aacaagtaca tcgtcagaga cccaaagtag 1440
 aagacaactg ggctcaggat tggttggcctc ccattgattg aaaccatgga ccagagaaac 1500
 atcaaggctg ttttggctac tcagttcaat gatttctctt tgggaaccag acacgatttc 1560
 ttgtactcct tggtgggtga cggtattttc accttggacg gtgctggctg gaaacatagt 1620
 agaactatgt tgagaccaca gtttgcctaga gaacaggttt ctacgtcaa gttgttggag 1680
 ccacacgttc aggtgttctt caagcacgtt agaaagcacc ggggtcaaac gttcgacatc 1740
 caagaattgt tcttcagggt gaccgtcgac tccgccaccg agttcttgtt tggtgagtct 1800
 gctgaatcct tgagggacga atctattgga ttgacctcaa ccaccaagga tttcgatggc 1860
 agaagagatt tcgctgacgc tttcaactat tcgcagactt accaggccta cagatttttg 1920
 ttgcaacaaa tgtactggat cttgaatggc tcggaattca gaaagtcgat tgctgtcgtg 1980
 cacaagtttg ctgaccacta tgtgcaaaaag gctttggagt tgaccgacga tgacttgcag 2040

ES 2 298 276 T3

	aaacaagacg gctatgtgtt cttgtacgag ttggctaagc aaaccagaga cccaaaggtc	2100
	ttgagagacc agttattgaa ctttttggtt gccggtagag acacgaccgc cggtttggtg	2160
5	tcatttgttt tctacgagtt gtcaagaaac cctgaggtgt ttgctaagtt gagagaggag	2220
	gtgaaaaca gatttggact cgggtgaagaa gctcgtgttg aagagatctc gtttgagtcc	2280
10	ttgaagtctt ctgagtactt gaaggctgtc atcaatgaaa ccttgagatt gtaccacatc	2340
	gttccacaca acttttagagt tgctaccaga aacactaccc tcccaagagg tgggtggtgaa	2400
15	gatggatact cggcaattgt cgtcaagaag ggtcaagttg tcatgtacac tgttattgct	2460
	accacagag ecccaagtat ctacggtgcc gacgctgacg tcttcagacc agaaagatgg	2520
	tttgaaccag aaactagaaa gttgggctgg gcatacgttc cattcaatgg tggccaaga	2580
20	atctgttttg ctcaacagtt tgccttgacc gaagcttcat acgtcactgt cagattgctc	2640
	caggagtttg cacacttgct tatggacca gacacgaat atccaccaa attgcagaac	2700
25	accttgacct tgcgctctt tgatggtgct gatgttagaa tgtactaagg ttgcttttcc	2760
	ttgctaattt tcttctgtat agcttggtga tttaaattga atcggcaatt gatttttctg	2820
30	ataccaataa ccgtagtgg atttgacca aaccgttcaa actttttgtt ctctcggtga	2880
	cgtgctcgt catcagcact gtttgaagac gaaagagaaa attttttgta aacaacactg	2940
	tccaaattta cccaacgtga accattatgc aaatgagcgg ccctttcaac tggtcgctgg	3000
35	aagcattcgg ggatatctac aacgccctta agtttgaac agacattgat ttagacacca	3060
	tagatttcag cggcatcaag aatgaccttg ccacatttt gagacccca acaccaactgg	3120
40	aagaatcacg ccagaaacta ggcgatggat ccaagcctgt gaccttgccc aatggagacg	3180
	aagtggagtt gaaccaagcg ttcttagaag ttaccacatt attgtcgaat gagtttgact	3240
45	tggaccaatt gaacgcggca gagttgttat actacgctgg cgacatatcc tacaagaagg	3300
	gcacatcaat cgcagacagt gccagattgt cttattattt gagagcaaac tacatcttga	3360
	acatacttgg gtatttgatt tcgaagcagc gattggattt gatagtcacg gacaacgacg	3420
50	cgttgtttga tagtattttg aaaagttttg aaaagatcta caagttgata agcgtgttga	3480
	acgatatgat tgacaagcaa aaggtgacaa gcgacatcaa cagtctagca ttcacatatt	3540
55	gcacaaacta ctcgagaggt caactattct ccgcacaaga acttttggga ctggttttgt	3600
	ttggattggt cgacatctat ttcaaccagt ttggcacatt agacaactac aagaaggat	3660
60	tggcattgat actgaagaac atcagcgatg aagacatott gatcatcac ttctcccat	3720
	cgacactaca attgtttaag ctggtgttgg acaagaaaga cgacgctgca gttgaacagt	3780
65	tctacaagta catcacttca acagtgtcac gagactacaa ctccaacatc ggctccacag	3840
	ccaaagatga tctcgatttg tccaaaacca aactcagtg ctttgagggtg ttgacgagtt	3900

ES 2 298 276 T3

<210> 11

<211> 3668

<212> ADN

5 <213> *Candida tropicalis*

<400> 11

10	cctgcagaat tcgcggccgc gtcgacagag tagcagttat gcaagcatgt gatttgtggt	60
	tttgcaacct gtttcacga caaatgatcg acagtcgatt acgtaatcca tattatttag	120
15	aggggtaata aaaaataaat ggcagccaga atttcaaaca ttttgcaaac aatgcaaaaag	180
	atgagaaact ccaacagaaa aaataaaaaa actccgcagc actccgaacc aacaaaacaa	240
	tggggggcgc cagaattatt gactattgtg actttttttt attttttccg ttaactttca	300
20	ttgcagtga gttgtttaca cggggtggtg atggtgttgg tttctacaat gcaagggcac	360
	agttgaaggt ttccacataa cgttgcacca tatcaactca atttatcctc attcatgtga	420
25	taaaagaaga gccaaaaggt aattggcaga ccccccaagg ggaacacgga gtagaaagca	480
	atggaaacac gcccatgaca gtgccattta gccacaaca catctagtat tttttttttt	540
	ttttgtgcgc aggtgcacac ctggacttta gttattgccc cataaagtta acaattctac	600
30	ctttggctct ccagtgctct ccgctccag atgctcggtt tacaccctcg agctaaccac	660
	aacacaacac ccattgagggg aatgggcaaa gttaaacact tttggtttca atgattccta	720
35	tttgctactc tttgttttg tgttttgatt tgcacctgt gaaataaacg acaattatat	780
	ataccttttc gtctgtctc caatgtctct tttgtctgc attttgcttt ttgcttttg	840
	cttttgcaact ctctcccaact ccacaaatca gtgcagcaac acacaagaa gaaaaataaa	900
40	aaaacctaca ctatgtctc ctctccatcg tttgtcagg aggttctcgc taaccactagt	960
	ccttacatcg agtactttct tgacaactac accagatggt actacttcat ccttttggtg	1020
45	cttttttctg tgaacttcat cagcttgctc cacacaaagt acttggaaag caggttccac	1080
	gccaaagcgc tcggtaacgt cgtgttggtat cctacgtttg gtatcgctac tccgttgatc	1140
50	ttgatctact taaagtcgaa aggtacagtc atgaagtttg cctggagctt ctggaacaac	1200
	aagtacattg tcaaagacc aaagtacaa accactggcc ttagaattgt cggcctccca	1260
	ttgattgaaa ccatagacc agagaacatc aaagctgtgt tggctactca gttcaacgat	1320
55	ttctccttgg gaactagaca cgatttcttg tactccttgt tgggcgatgg tatttttacc	1380

ES 2 298 276 T3

ttggacggtg ctggctggaa acacagtaga actatgttga gaccacagtt tgctagagaa 1440
 5 caggtttccc acgtcaagtt gttggaacca cacgttcagg tgttttcaa gcacgttaga 1500
 aaacaccgcg gtcagacttt tgacatccaa gaattgttct tcagattgac cgtcgactcc 1560
 gccaccgagt tcttgtttgg tgagttctgt gaatccttga gagacgactc tgttggtttg 1620
 10 accccaacca ccaaggattt cgaaggcaga ggagatttcg ctgacgcttt caactactcg 1680
 cagacttacc aggcctacag atttttgttg caacaaatgt actcgatttt gaatggcgcg 1740
 15 gaattcagaa agtcgattgc catcgtgcac aagtttctg accactatgt gcaaaaggct 1800
 ttggagttga ccgacgatga cttgcagaaa caagacggct atgtgttctt gtacgagttg 1860
 20 gctaagcaaa ctagagaccc aaaggtcttg agagaccagt tgttgaacat tttggttgcc 1920
 ggtagagaca cgaccgccg tttgttctcg tttgttctt acggttctc gagaaacctt 1980
 gaagtgtttg ccaagttgag agaggaggtg gaaaacagat ttgactcgg cgaagaggct 2040
 25 cgtgttgaaq agatctcttt tgagtccttg aagtctctg actacttga ggctgtctc 2100
 aatgaagcct tgagattgta cccatctgtt ccacacaact tcagagttgc caccagaaac 2160
 30 actacccttc caagaggcgg tggtaaagac ggtgtctgc caattgttgt caagaagggt 2220
 caagttgtca tgtacactgt cattggtacc cacagagacc caagtatcta cggtgccgac 2280
 35 gccgacgtct tcagaccaga aagatggttc gagccagaaa ctgaaaagt gggctgggca 2340
 tatgttccat tcaatggtgg tccaagaatc tgtttgggtc agcagtttgc cttgactgaa 2400
 gcttcatacg tcaactgtcag attgtcccaa gagtgtgaa acgtgtccct ggatccaaac 2460
 gctgagtacc caccaaaatt gcagaacacc ttgacctgtt cactctttga tgggtgtgac 2520
 45 gttagaatgt tctaagggtg cttatccttg ctagtgttat tctagtttg tgtattttaa 2580
 ttgaatcggc gattgatatt tctggtacta ataactgtag tgggttttga ccaaaaccgt 2640
 tcaaaccttt tttttttttt tcttccccct acctcggttg ctggtctcgc agcactgttt 2700
 50 gaaaacgaaa aaagaaaatt ttttgtaaac aacattgcc aaacttacc aacgtgaacc 2760
 attataacca aatgagcggc gctttcaact ggtcactgga gctattcggg gatattctaca 2820
 55 acacccttaa gtttgaggaa gacattgatt tagacaccat agatttcage ggcattcaaga 2880
 atgacctgtt ccacattttg acaaccccaa caccactgga agtatcgcgc cagaaactag 2940
 gcgatggatc caagcctgtg gccttgccca atggagacga agtgaggttg aaccaagcgt 3000
 60 tcttagaagt taccacatta ttgtcgaacg agtttgactt ggaccaattg aacgcggccg 3060

ES 2 298 276 T3

agttgttata ctacgcgggc gacatatcct acaagaaggc cacatcaatt gccgacagtg 3120
 ccagattgtc ttactatttg agagcaaaact acatcttgaa cataacttggg tacttttattt 3180
 5 cgaagcagcg attggatgtg atagtcaccg acaacaacgc gttgtttgat aatattttga 3240
 aaagtittga aaagatctac aagttgataa gcgcggttgaa cgatatcatt gacaagcaaa 3300
 10 aggtgacaag cgacatcaac agtctagcat ttatcaactg catcaactac tcgaggggtc 3360
 aactattctc cgcacaagaa cttctgggac tggttttggt tggatttggt gacaactatt 3420
 tcaaccagtt tggctcatta gacaactaca agaaagtatt ggcattcata ctgaagaaca 3480
 15 tcagtgtgga agatatcttg atcgtacgtt tctcccatc gacaactaca ttgtttaagc 3540
 tgggtgttga taagaaagac gacgccactg ttgaccagtt ctacaagtac atcaccctca 3600
 20 cagtgtcgca agactacaac tccaacatcg gagccacagc caaagatgat atcgatttgt 3660
 ccaaagcc 3668

25 <210> 12
 <211> 3348
 <212> ADN
 30 <213> *Candida tropicalis*
 <400> 12

35 gatgtggtgc ttgatttctc gagacacatc cttctgaggt gccatgaatc tgtacctgtc 60
 tctaagcaca gggaaactgt tcaacacett attgcataat ctgtctattg caagcgtgtg 120
 ctgcaacgat atctgccaag gtatatagca gaacgtgctg atggttcttc eggctatatt 180
 40 ctgttggtag ttctgcaggt aaatttggtg gtcaggtagt ggagggaggt ttgtatcggt 240
 tgtgttttct tcttctcttc tctctgatc aacctccacg tctcttctcg gttctgtgtc 300
 45 tgtgtctgag tctactgtt ggattaagtc catcgcatgt gtgaaaaaaa gtagcgctta 360
 tttagacaac cagttcgttg ggcgggtatc agaaatagtc tgttgtgcac gaccatgagt 420
 atgcaacttg acgagacgtc gttaggaatc cacagaatga tagcaggagc cttactacgt 480
 50 gagagattct gcttagagga tgttctcttc ttgttgattc cattaggttg gtatcatctc 540
 cgggtggtgac aacttgacac aagcagttcc gagaaccacc cacaacaatc accattccag 600
 55 ctatcacttc tacatgtcaa cctacgatgt atctcatcac catctagttt cttggcaatc 660
 gtttatttgt tatgggtcaa catccaatc aactccacca atgaagaaga aaaacggaaa 720
 60 gcagaatacc agaatgacag tgtgagttcc tgaccattgc taatctatgg ctatatctag 780
 tttgctatcg tgggatgtga tctgtgtcgt cttcatttgc gtttgtgtt atttcgggta 840

65

ES 2 298 276 T3

tgaatattgt tataactaaat acttgatgca caaacatggc gctcgagaaa tcgagaatgt 900
 gatcaacgat ggggttctttg gggttcogett acctttgcta ctcatgagag ccagcaatga 960
 5 gggccgactt atcgagttca gtgtcaagag attcgagtcg gcgccacatc cacagaacaa 1020
 gacattggtc aaccgggcat tgagcgttcc tgtgatactc accaaggacc cagtgaatat 1080
 10 caaagcgatg ctatcgaccc agtttgatga cttttccctt gggttgagac tacaccagtt 1140
 tgcgccggtt ttggggaaaag gcatctttac tttggacggc ccagagtggg agcagagccg 1200
 15 atctatgttg cgtccgcaat ttgccaaga tggggtttct catatcctgg atctagaacc 1260
 gcattttgtg ttgcttcgga agcacattga tggccacaat ggagactact tcgacatcca 1320
 ggagctctac ttccggttct cgatggatgt ggcgacgggg tttttgttg gcgagtctgt 1380
 20 ggggtcgttg aaagacgaag atgcgaggtt cctggaagca ttcaatgagt cgcagaagta 1440
 tttggcaact agggcaacgt tgcacgagtt gtactttctt tgtgacgggt ttaggtttcg 1500
 25 ccagtacaac aaggttggtc gaaagttctg cagccagtgt gtccacaagg cgtagatgt 1560
 tgcaccggaa gacaccagcg agtacgtgtt tctccgag ttggtcaaac acactcgaga 1620
 30 tcccgttgtt ttacaagacc aagcgttgaa cgtcttgctt gctggacgag acaccaccgc 1680
 gtcgttatta tcgtttgcaa catttgagct agcccggaat gaccacatgt ggaggaagct 1740
 acgagaggag gttatctga cgatgggacc gtccagtgt gaaataaccg tggccgggtt 1800
 35 gaagagttgc cgttacctca aagcaatcct aaacgaaact cttegactat acccaagtgt 1860
 gcctaggaac gcgagatttg ctacgaggaa tacgacgctt cctcgtggcg gaggtccaga 1920
 40 tggatcgttt ccgattttga taagaaaggg ccagccagtg gggatattca tttgtgctac 1980
 acacttgaat gagaaggtat atgggaatga tagccatgt ttcgaccgg agagatgggc 2040
 45 tgcgttagag ggcaagagtt tgggctggtc gtatcttcca ttcaacggcg gcccgagaag 2100
 ctgcccttgg cagcagtttg caatccttga agcttcgtat gttttggctc gattgacaca 2160
 50 gtgctacacg acgatacagc ttagaactac cgagtacca ccaaagaaac tcgttcactc 2220
 cacgatgagt cttctcaacg ggggtgtacat ccgaactaga acttgattat gtgtttatgg 2280
 ttaatcgggg caaagcactg caagtcattg atgtttgtgg aagcccagca ttggtgttcc 2340
 55 ggagcatcaa taaccaatgt cttgaagggt ttgattttct tgaccttctt ctctctgagc 2400
 ttctttccgt caaacttgta cagaatggcc atcatttcag gaacaaccac gtacgacggc 2460
 60 cggtagcgca tctggagtat ctccgctcg ttcaagtagc acgaaaacag caacgacgct 2520
 accatctgct tcccaatctt gacaccacaa gataccctg cggcttcag gatcaaaaac 2580

65

ES 2 298 276 T3

gtcggcaacc cccggtatat gtccatgtaa ttctccatgg ccacctccat caacacactg 2640
 atggagcgac tgacgggtgcc accactgccc tcgggtgagt caaggcagta tgatgccggg 2700
 5 atccagtact ccaatgggaa cctctgcacg gtgtcgtcgc agtttttgag gcgtatttcg 2760
 atccatgac gttcttttgt gctgtagtat aacgagctct tgggtgtcctt gaaatggaac 2820
 10 aggttggtatg tgttgttgag ttgtctgcg tgcttggtt gcaagtcttc gatcgagcgt 2880
 agtgagtaga cagttggcgg ggggtgtggc tcgggcttta ttctgtgttt gtgtttcctt 2940
 cttagtcttg gaatgacgt gttatcgacg gttcgtagta taagtagcgc caatatgaga 3000
 15 atgtatatcc gcatcaccac agactcttca gcctgttaca acgactgagg ctgttgcccg 3060
 tgtgaccaat tggttctttt ggtgacctag attggctccg cagggaaagc aagggtcgt 3120
 20 aggggggcat accaaacaag gtctgtaat cagtatctat ggtgctacca tgtgtgtggt 3180
 tggggggaaa ttccgcatt ttgtgtaac gaaagtctta gaaagtcttc gtgggtctcg 3240
 agaatctgct ggaaccatcc acccgattt cegtgcaca agtgggaaga gcaatcaacc 3300
 25 caccctgctt tgccaatca gccattcccc tgggaatata aattcaac 3348

30 <210> 13
 <211> 3826
 <212> ADN
 <213> *Candida tropicalis*
 35 <400> 13

tggaatcgcc agacttgctc acttttgact ccttcgaaa ctcaaatgac gttcaggcgg 60
 40 tgctcaacga aacgcctcgt atctaccggg gggtaaccag aaacatgaag acagctacgt 120
 gcaacaacac gttagcagc ggaggaggca aagacggcaa ggaacctatc ttgtgcaga 180
 45 agggacagtc cgttgggttg attactattg ccacgcagac ggaccagag tattttgggg 240
 ccgacgctgg tgagtttaag ccgagagat ggtttgattc aagcatgaag aacttggggt 300
 gtaactactt gccgttcaat gctgggccc ggacttgctt ggggcagcag tacactttga 360
 50 ttgaagcgag ctacttgeta gtccggttg ccagaccta ccgggcaata gatttgcagc 420
 caggatcgcc gtaccaccca agaaagaagt cgttgatcaa catgagtgt gccgacgggg 480
 55 tgtttgtaaa gctttataag gatgtaacgg tagatggata gttgtgtagg aggagcggag 540
 ataaattaga ttgattttg tgtaaggttt tggatgtcaa cctactcgc acttcatgca 600
 gtgtgtgtga cacaaggggt tactacgtgt gcgtgtgcgc caagagacag cccaaggggg 660
 60 tggtagtgtg tgttggcggg agtgcagtgt acacaacgcg tgggttcttg ccaatgggtg 720

65

ES 2 298 276 T3

	actaagtgca ggtaagcagc gacctgaaac attcctcaac gettaagaca ctgggtggtag	780
5	agatgcggac caggctattc ttgtcgtgct acccggcgca tggaaaatca actgcgggaa	840
	gaataaattt atccgtagaa tccacagagc ggataaattt gcccacctcc atcatcaacc	900
	acgccgccac taactacatc actccctat tttctctctc tctctttgtc ttaactccgt	960
10	cccgtttcct tagccacaga tacacacca ctgcaaacag cagcaacaat tataaagata	1020
	cgccaggccc accttctttc tttttcttca cttttttgac tgcaactttc tacaatccac	1080
15	cacagccacc accacagccg ctatgattga acaactccta gaatattggt atgtcgttgt	1140
	gccagtgttg tacatcatca aacaactcct tgcatacaca aagactcggc tcttgatgaa	1200
20	aaagttaggt gctgctccag tcacaaacaa gttgtacgac aacgctttcg gtatcgtcaa	1260
	tggatggaag gctctccagt tcaagaaaga gggcagggtc caagagtaca acgattacaa	1320
	gtttgaccac tccaagaacc caagcgtggg cacctacgtc agtattcttt tcggcaccag	1380
25	gatcgtcgtg accaaagatc cagagaatat caaagctatt ttggcaacc agtttggtga	1440
	tttttctttg ggcaagaggc acactctttt taagcctttg ttaggtgatg ggatcttcac	1500
30	attggacggc gaaggctgga agcacagcag agccatgttg agaccacagt ttgccagaga	1560
	acaagttgct catgtgacgt cgttggaacc acacttccag ttgttgaaga agcatattct	1620
35	taagcacaag ggtgaatact ttgatatcca ggaattgttc tttagattta ccgttgattc	1680
	ggccacggag ttcttatttg gtgagtcctg gcactcctta aaggacgaat ctattggtat	1740
	caaccaagac gatatagatt ttgctggtag aaaggacttt gctgagtcgt tcaacaaagc	1800
40	ccaggaatac ttggctatta gaaccttggt gcagacgttc tactggttgg tcaacaacaa	1860
	ggagttttaga gactgtacca agctggtgca caagttcacc aactactatg ttcagaaagc	1920
45	tttggtgct agcccagaag agcttgaaaa gcaaagtggg tatgtgttct tgtacgagct	1980
	tgtcaagcag acaagagacc ccaatgtgtt gcgtgaccag tctttgaaca tcttggtggc	2040
50	cggaaagagc accactgctg ggttggtgtc gtttgctgtc tttgagttgg ccagaccccc	2100
	agagatctgg gccaaagttg gagaggaaat tgaacaacag tttggtcttg gagaagactc	2160
	tcgtgttgaa gagattacct ttgagagctt gaagagatgt gagtacttga aagcgttcc	2220
55	taatgaaacc ttgcgtattt acccaagtgt cccaagaaac ttcagaatcg ccaccaagaa	2280
	cacgacattg ccaagggggc gtggttcaga cggtagctcg ccaatcttga tccaaaagg	2340
60	agaagctgtg tcgtatggtg tcaactctac tcatttggac cctgtctatt acggccctga	2400

65

ES 2 298 276 T3

5 tgcctgctgag ttcagaccag agagatgggt tgagccatca accaaaaagc tcggctgggc 2460
 ttacttgcca ttcaacggtg gtccaagaat ctggttgggt cagcagtttg ccttgacgga 2520
 10 agctggctat gtgttggtta gattggtgca agagtctctc cacgttaggc tggaccaga 2580
 cgaggtgtac ccgccaaaga ggttgaccaa cttgaccatg tggttgacag atggtgetat 2640
 15 tgtcaagitt gactagcggc gtggtgaatg cgtttgattt tgtagttctt gtttgacgta 2700
 atgagataac tattcagata agcgcagtggt atgtacgttt tgtaagagtt tctttacaac 2760
 20 cttggtgggg tgtgtgaggt tgaggttgca tcttggggag attacacctt ttgcagctct 2820
 ccgtatacac ttgtactctt tgtaacctct atcaatcatg tggggggggg ggttcattgt 2880
 ttggccatgg tgggtgcatgt taaatccgcc aactaccaa tctacatga aactcaagca 2940
 25 caetaaaaaa aaaaaagatg ttgggggaaa actttggttt ccottcttag taattaaaca 3000
 ctctcactct cactctcact ctctccactc agacaaacca accacctggg ctgcagacaa 3060
 ccagaaaaaa aaagaacaaa atccagatag aaaaacaaag ggctggacaa ccataaataa 3120
 30 acaatctagg gtctactcca tcttccactg tttctctctt ttcagactta gctaacaaac 3180
 aactcacttc accatggatt acgcaggcat cagcgtggc tocatcagag gcgaggcctt 3240
 35 gaagaaactc gcagaattga ccacccagaa ccagccatcc agcttgaaaq aaatcaacac 3300
 cggcatccag aaggacgact ttgccaagtt gttgtctgcc acccggaaaa tccccaccas 3360
 40 gcacaagttg aacggcaacc acgaattgtc tgaggtcgcc attgccaaaa aggagtacga 3420
 ggtgttgatt gccttgagcg acgccacaaa agaccaatc aaagtgcctt cccagatcaa 3480
 gatcttgatt gacaagttca aggtgtactt gtttgagttg cctgaccaga agttctctta 3540
 45 ctccatcgtg tccaactcgg tcaacatcgc ccctggacc ttgctcgggg agaagttgac 3600
 cacgggcttg atcaacttgg ccttcagaa caacaagcag cacttgacg aggtcattga 3660
 50 catcttcaac gagtccatcg acaagttctt tggcaacacg gagccgcaat tgaccaactt 3720
 cttgaccttg tgcggtgtgt tggacgggtt gattgacct gccaaattct tgagcgtgtc 3780
 ctgcgggacc tcaagatct tcttgaactt ggactcgtat gtggac 3826

<210> 14

<211> 3910

<212> ADN

<213> *Candida tropicalis*

ES 2 298 276 T3

<400> 14

	ttacaatcat ggagctcgct aggaaccag atgtctggga gaagctccgc gaagaggtca	60
5	acacgaactt tggcatggag tcgccagact tgctcacttt tgactctctt agaagctcaa	120
	agtacgttca ggcggtgctc aacgaaacgc ttcgtatcta cccgggggtg ccacgaaaca	180
10	tgaagacagc tacgtgcaac acgacgttgc cgcgtggagg aggcaaagac ggtaaggaac	240
	ctattttggg gcagaagggc cagtcocgtt gggtgattac tattgccacg cagacggacc	300
15	cagagtattt tggggcagat gctggtgagt tcaaaccgga gagatggtt gattcaagca	360
	tgaagaactt ggggtgtaag tacttgccgt tcaatgctgg gccccggact tgtttggggc	420
	agcagtacac tttgattgaa gcgagctatt tgctagtcag gttggcgag acctaccggg	480
20	taatcgattt gctgccaggg tcggcgtaac caccaagaaa gaagtcgttg atcaatatga	540
	gtgctgccga tggggtggtt gtaaagtttc acaaggatct agatggatat gtaagggtgtg	600
25	taggaggagc ggagataaat tagatttgat tttgtgtaag gtttagcacg tcaagctact	660
	ccgcactttg tgtgtaggga gcacatactc cgtctgcgcc tgtgccaaqa gaaggcccag	720
30	gggtagtgtg tgggtgtgga agtgcatgtg acacaatacc ctggttcttg ccaattgggg	780
	atttagtgta ggtaagctgc gacctgaaac actcctcaac gcttgagaca ctgggtggga	840
35	gagatgcggg ccaggaggct attcttgctg tgctaccggt gcacggaaaa tcgattgagg	900
	gaagaacaaa tttatccgtg aaatccacag agcggataaa tttgtcacat tgcctcgttg	960
	cccaccaca gcattctctt ttctctctct ttgtcttact ccgtcctgt ttctttatcc	1020
40	agaaatacac accaactcat ataaagatac gctagcccag ctgtctttct tttcttcac	1080
	tttttttggg gtgttgcttt tttggctgct actttctaca accaccacca ccaccaccac	1140
45	catgattgaa caaatccctag aatattggta tattgttgtg cctgtgttgt acatcatcaa	1200
	acaactcatt gcctacagca agactcgcgt cttgatgaaa cagttgggtg ctgctccaat	1260
50	cacaaaccag ttgtacgaca acgttttcgg tatcgtcaac ggatggaagg ctctccagtt	1320
	caagaaagag ggcagagctc aagagtacaa cgatcacaaag tttgacagct ccaagaaccc	1380
	aagcgtcggc acctatgtca gtattctttt tggcaccaag attgtcgtga ccaaggatcc	1440
55	agagaatata aaagctattt tggcaaccca gtttgcgat tttcttttg gcaagagaca	1500
	cgtctttttt aaacctttgt taggtgatgg gatcttcacc ttggacggcg aaggctggaa	1560
60	gcatagcaga tccatgttaa gaccacagtt tgccagagaa caagttgctc atgtgacgtc	1620
	gttggaaacca cacttcaggt tgttgaagaa gcatatcctt aaacacaagg gtgagtactt	1680
65	tgatatccag gaattgttct ttagatttac tgctgactcg gccacggagt tcttatttgg	1740
	tgagtcctgt cactccttaa aggacgaaac tatcggatc aaccaagacg atatatagattt	1800

ES 2 298 276 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

tgctggtaga aaggactttg ctgagtcgtt caacaaagcc caggagtatt tgtctattag 1860
 aattttggtg cagaccttct actggttgat caacaacaag gagtttagag actgtaccaa 1920
 gctggtgcac aagtttacca actactatgt tcagaaagct ttggatgcta cccagagga 1980
 acttgaaaag caaggcgggt atgtgttctt gtatgagctt gtcaagcaga cgagagaccc 2040
 caagggtgtg cgtgaccagt ctttgaacat cttgttgga ggaagagaca ccaactgctg 2100
 gttgtgtgct tttgctgtgt ttgagttggc cagaaacca cactctgtgg ccaagttgag 2160
 agaggaaatt gaacagcagt ttggtcttgg agaagactct cgtgttgaag agattacctt 2220
 tgagagcttg aagagatgtg agtacttgaa agcgttccct aacgaacct tgcgtgttta 2280
 cccaagtgtc ccaagaaact tcagaatgc caccaagaat acaacattgc caaggggtgg 2340
 tggccagac ggtaccagc caatcttgat ccaaaagga gaagggtgtg cgtatggtat 2400
 caactctacc cacttagatc ctgtctatta tggccctgat gctgctgagt tcagaccaga 2460
 gagatggttt gagccatcaa ccagaaagct cggctgggct tacttgccat tcaacggtgg 2520
 gccacgaatc tgtttgggtc agcagtttgc cttgaccgaa gctggttacg ttttggtcag 2580
 attggtgcaa gagttctccc acattaggct ggaccagat gaagtgtatc caccaaagag 2640
 gttgaccaac ttgaccatgt gtttgcagga tgggtgctatt gtcaagttt actagtacgt 2700
 atgagtgcgt ttgattttgt agtttctgtt tgcagtaatg agataactat tcagataagg 2760
 cgggtggatg tacgttttgt aagagtttcc ttacaacct ggtgggtgtg tgaggttgca 2820
 tcttagggag agatagcacc ttttgcagct ctccgtatac agttttactc tttgtaacct 2880
 atgccaatca tgtggggatt cattgtttgc ccattggtgt gcattgcaaaa tcccccaac 2940
 tacccaatct cactgaaac tcaagcacac tagaaaaaaa agatgttgcg tgggttcttt 3000
 tgatgttggg gaaaactttc gtttctttc tcagtaatta aacgttctca ctacagacaaa 3060
 ccacctgggc tgcagacaac cagaaaaaac aaatccaga tagaagaaga aaggcctgga 3120
 caaccataaa taaacaacct agggctccact ccattcttca cttcttcttc ttcagactta 3180
 tctaacaaac gactcacttc accatggatt acgcaggtat cacgcgtggg tccattcagag 3240
 gcgaagcctt gaagaaactc gccgagttga ccattccagaa ccagccatcc agcttgaaag 3300
 aaatcaacac cggcatccag aaggacgact ttgccaaagt gttgtcttcc accccgaaaa 3360
 tccacaccaa gcacaagttg aatggcaacc acgaattgtc cgaagtcgcc attgccaaaa 3420
 aggagtacga ggtgttgatt gccttgagcg acgccacgaa agaaccaatc aaagtcacct 3480

ES 2 298 276 T3

	cccagatcaa gatcttgatt gacaagttca aggtgtact: gtttgagttg cccgaccaga	3540
	agttctccta ctccacgtg tccaactcg ttaacattgc cccctggacc ttgctcggtg	3600
5	agaagttgac caagggttg atcaacttgg cgttcacaga caacaagcag caettggacg	3660
	aagtcacga catcttcaac gagttcatcg acaagttctt tggcaacaca gagccgcaat	3720
10	tgaccaactt cttgaccttg tccggtgtgt tggacgggtt gattgaccat gccaaacttc	3780
	tgagcgtgtc ctccaggacc tccaagatct tcttgaactt ggactcgttc gtggaceact	3840
	cggacttctt gaacgacgtg gagaactact ccgacttttt gtacgacgag ccgaaccggt	3900
15	accacaactt	3910

<210> 15

20 $\langle 211 \rangle$ 3150

<212> ADN

<213> *Candida tropicalis*

25 <400> 15

	gaattctttg gatctaattc cagctgatct tgcataatct tatcaacgta gttgtgatca	60
30	ttgtttgtct gaattataca caccagtggga agaataatggt ccaatttgca cgtcccactg	120
	gcatttgttg tttgtggggg ggggggggtg cacacatttt tagtgccatt ctttgttgat	180
35	tacccctccc cctatcatt cattcccaca ggattagtgt tttctcact ggaattcgct	240
	gtccacctgt caaccccccc ccccccccc cccactgccc taccctgccc tgccttgac	300
	gtcctgtgtt ttgtgctgtg tctttcccac gctataaaag ccttgccgtc cggccaaggt	360
40	ttttccaccc agccaaaaaa acagtctaaa aaatttggt gatccttttt ggttgcaagg	420
	ttttccacca ccacttccac caccctcaact attcgaacaa aagatgctcg atcagatctt	480
45	acattactgq tacattgtct tgcatttgtt ggcattatc aaccagatcg tggctcatgt	540
	caggaccaat tatttgatga agaaattggg tgcataagcca ttcacacacg tccaacgtga	600
	cgggtgggtg ggcttcaaat tgggcgtga attcctcaaa gcaaaaagtg ctgggagact	660
50	ggttgattta atcactctcc gtttccacga taatgaggac actttctcca gctatgcttt	720
	tggcaaccat gtggtgttca ccagggaacc cgagaatata aaggcgcttt tggcaaccca	780
55	gtttggtgat ttttcattgg gcagcagggg caagttcttc aaaccattat tggggtacgg	840
	tatcttcaca ttggacgccc aaggctggaa gcacagcaga gccatgttga gaccacagtt	900
	tgccagagaa caagttgctc atgtgacgtc gttggaacca cacttccact tgttgaagaa	960
60	gcatactctt aaacacaagg gtgagtactt tgatatccag gaattgttct ttagatttac	1020

ES 2 298 276 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

tgtcgactcg gccacggagt tcttatttgg tgagtccttg cactccctaa aggacgagga 1080
 aattggctac gacacgaaag acatgtctga agaaagacgc agatttgccg acgcgttcaa 1140
 caagtcgcaa gtctacgtgg ccaccagagt tgctttacag aacttgtaact ggttggtcaa 1200
 caacaaagag ttcaaggagt gcaatgacat tgtccacaag tttaccaact actatgttca 1260
 gaaagccttg gatgctaccc cagaggaact tgaaaagcaa ggcgggtatg tgttcttgta 1320
 tgagcttgtc aagcagacga gagaccccaa ggtgttgctg gaccagtctt tgaacatctt 1380
 gttggcagga agagacacca ctgctgggtt gttgtccttt gctgtgtttg agttggccag 1440
 aaaccacac atctgggcca agttgagaga ggaaattgaa cagcagtttg gtcttgagga 1500
 agactctcgt gttgaagaga ttacctttga gagcttgaag agatgtgagt acttgaaggc 1560
 cgtgttgaaac gaaactttga gattacaccc aagtgtccca agaaacgcaa gatttgcat 1620
 taaagacacg actttaccaa gaggcggtgg cccaacggc aaggatccta tcttgatcag 1680
 gaaggatgag gtggtgcagt actccatctc ggcaactcag acaaactcctg cttattatgg 1740
 cgcgatgct gctgatttta gaccggaaag atgggttgaa ccatcaacta gaaacttggg 1800
 atgggctttc ttgccattca acggtggtcc aagaatctgt ttgggacaac agtttgcttt 1860
 gactgaagcc ggttacgttt tggtagact tgttcaggag ttccaaact tgtcacaaga 1920
 cccgaaacc aagtaccac cactagatt ggcacacttg acgatgtgct tgtttgacgg 1980
 tgcacacgtc aagatgtcat aggtttcccc atacaagtag ttcagtaatt atacactgtt 2040
 tttactttct cttcatacca aatggacaaa agttttaagc atgcctaaca acgtgaccgg 2100
 acaattgtgt cgactagta tgtaacsaatt gtaaaaatag tgtacactaa tttgtggtgg 2160
 ccggagataa attacagttt ggttttgtgt aaactcgcg atatctctgg cagtttctct 2220
 tctccgcagc agctttgccg cgggtttgct ctggggccaa caaattcaaa agggggagaa 2280
 acttaacacc ccttatctct ccaactctagg ttgtagctct tgtggggatg caattgtcgt 2340
 acgtttttta tgttttgtct agactttgat gattacgttg gatttcttat gtctgaggcg 2400
 tgcttgaaag aagtgtcaaa atgtgacagg cgacgtatatt cgacatgaac gcgaaagggt 2460
 tatttgcatc aatacgaggg gctgactcta gtctaggatg gcagtcctag gttgcaaaca 2520
 tgttgacca tatectctct ggagttggtc gacctcgctt acgccacctt cagegatcgg 2580
 cactttccgt tgttcaatat ttctccttcc cattgttcca ggggttatca acaacgttgc 2640
 cggcctctct cccaaattac aagaaaaata aattgtcgca cggcaccgat ctgtcaaaga 2700
 tacagataaa ccttaaatct gcaaaaacaa gacccctccc catagcctag aagcaccagc 2760

ES 2 298 276 T3

5 aagatgatgg agcaactcct ccagtactgg tacatcgac tetctgtatg gttcatcctt 2820
 cgctactttg cttccacgc acgagccgtc tacttgccc acaagctcgg cgcggcgcca 2880
 10 ttcacgcaca ccagtacga cggctggtat gggttcaagt ttgggcggga gtttctcaag 2940
 gcgaagaaga tcgggcggca gacggacttg gtgcatgcgc ggttccgttg cggcatggac 3000
 accttctcga gctacacitt cggcatccat atcatcctta cccgggacct ggagaacatc 3060
 aagggcgtct tggcgacgca gttcgatgac ttctcgctcg gtggcaggat caggttcttg 3120
 aagccgttgt tggggtatgg gatattcagc 3150

15 <210> 16
 <211> 3579
 <212> ADN
 20 <213> *Candida tropicalis*
 <400> 16

25 aaaaccgata caagaagaag acagtcaaca agaacgttaa tgtcaaccag gcgccaagaa 60
 gacggttttg cggacttggg agaattgtgg atttgcccat gatgtttatg ttctggagag 120
 gtttttcaag gaatcgtcat cctccggcac cacaagaacc accagttaac gagatccata 180
 30 ttcacaaccc accgcaaggt gacaatgctc aacaacaaca gcaacaacaa caacccccac 240
 aagaacagtg gaataatgcc agtcaacaaa gagtggtagc agacgagggg gaacacgcaa 300
 gcaacagtg ttctgctgca agatcagcta cccgcttca tcaggaaaag caggagctcc 360
 35 caccaccata tgcctctcac ggcacacacc agcaggttag tctatagtag tctgtagtta 420
 agtcaatgca atgtaccaat aagaactatc cttcttaca ccaagttttc tgcgcgcct 480
 40 gtctggcaac agatgctggc cgacacactt tcaactgagt ttgggtctaga attcttgca 540
 atgcacgaca aggaactct tacaagaca acacttgctc tctgatgcca ctgcatcttg 600
 ctaagcetta tcaacgtaat tgagatcatt gttgtctga attatacaca ccagtggag 660
 45 aatcggctct aatctgcacg cctcatgggc attgtgtgtt ttgggggggg ggggggggt 720
 gcacacattt ttagtgcgaa tgtttgtttg ctggttccc ctcctccctc cccctatca 780
 50 tgcctcacagg attagttttt tctcactgg aatcgtctgt ccacctgtca acccctcac 840
 tgcctgccc tgcctgcac gccctgtgtt ttgtgcgtg gcactccac gctataaaag 900
 cctggcgta cggccaaggt tttcctcac agccaaaaa aaatttggct gatccttttg 960
 55 ggtgcaagg tttttacca ccaccaccac caccacctca actattcaaa caaggatgc 1020
 tcgaccagat cttccattac tggtagattg tcttgccatt gttggcatt atcaagcaga 1080

60

65

ES 2 298 276 T3

	tcgtggctca	tgccaggacc	aattatttga	tgaagaagtt	gggcgctaag	ccattcacac	1140
	atgtccaact	agacgggtgg	tttggcttca	aatttggtcg	tgaattcctc	aaagctaaaa	1200
5	gtgctgggag	gcaggttgat	ttaatcatct	cccgtttcca	cgataatgag	gacactttct	1260
	ccagctatgc	ttttggcaac	catgtggtgt	tcaccagggg	ccccgagaat	atcaaggcgc	1320
10	ttttggcaac	ccagtttggg	gatttttcat	tgggaagcag	ggtcaaattc	ttcaaaccat	1380
	tgttggggta	cggatatctc	accttggacg	gcgaaggctg	gaagcacagc	agagccatgt	1440
15	tgagaccaca	gtttgccaga	gagcaagttg	ctcatgtgac	gtcgttggaa	ccacatttcc	1500
	agttgttgaa	gaagcatatt	cttaagcaca	aggggtgaata	ctttgatatc	caggaattgt	1560
	tctttagatt	taccgttgat	tcagcgacgg	agttcttatt	tggtagtcc	gtgcactcct	1620
20	taagggacga	ggaaattggc	tacgatacga	aggacatggc	tgaagaaaga	cgcaaatttg	1680
	ccgacgcgtt	caacaagtgc	caagtctatt	tgtccaccag	agttgcttta	cagacattgt	1740
25	actggttggg	caacaacaaa	gagttcaagg	agtccaacga	cattgtccac	aagttcacca	1800
	actactatgt	tcagaaagcc	ttggatgcta	ccccagagga	acttgaaaaa	caaggcgggt	1860
30	atgtgttctt	gtaccagctt	gccaaacaga	cgaagagccc	caatgtgttg	cgtgaccagt	1920
	ctttgaacat	cttgttggct	ggaagggaca	ccactgctgg	gttgttgtcc	tttgcctgtg	1980
35	ttgagttggc	caggaaccca	cacatctggg	ccaagttgag	agaggaaatt	gaatcacact	2040
	ttgggctggg	tgaggactct	cgtgttgaag	agattacctt	tgagagcttg	aagagatgtg	2100
	agtacttgaa	agccgtgttg	aacgaaacgt	tgagattaca	cccaagtgtc	ccaagaaacg	2160
40	caagatttgc	gattaaagac	acgactttac	caagaggcgg	tggccccaac	ggcaaggatc	2220
	ctatcttgat	cagaaagaat	gaggtgggtc	aatactccat	ctcggcaact	cagacaaatc	2280
45	ctgcttatta	tggcgccgat	gctgctgatt	ttagaccgga	aagatggttt	gagccatcaa	2340
	ctagaaactt	gggatgggct	tacttgccat	tcaacggtgg	tccaagaatc	tgcttgggac	2400
50	aacagtttgc	tttgaccgaa	gccggttacg	ttttgggttag	acttgttcag	gaattcccta	2460
	gcttgtcaca	ggaccccgaa	actgagtacc	caccacctag	attggcacac	ttgacgatgt	2520
	gcttgtttga	cggggcatac	gtcaagatgc	aataggtttt	ggtttgactt	tgtttccata	2580
55	tgcaagtagt	tcagtaatta	cacactaatt	tgtggtggcc	ggcgataaat	taccgtttgg	2640
	ttttgtgtaa	aaattcggac	atctctgggt	gtttcccttc	tccgcagcag	ctttgccacg	2700
60	ggtttgctct	gcggccaaca	aattcgaaag	gggggggggg	gggggagaaa	gttaacaccc	2760

65

ES 2 298 276 T3

5 cctgttccca ccgtaggctg tagctcttgt ggggggatgt aattgtcgtc cgttttcattg 2820
 tttggcccaag actttgatga ttacgtaggc tttcttatgt ctaaggcgtg cttgacacaa 2880
 gtgtcaaaag gtgacaggcg acgttattcg acatgaacgc aaaagggtaa tttgcatcga 2940
 taagaggggt tgctcttgtt ctaagaagga cccccagggt tgcaaactatg ttgcactqca 3000
 10 tcccactcag agttgggtcga ccacgcctac gcttaccctc agcgatcggc actttccgtt 3060
 gctcaatatt totctccccc ctgcttcccc ccattgttcc agggattatc aacaacgttg 3120
 ccggtctcct ctcccccccc tccccccagt tatgtacaag aaaatttaat tgtcgcacgg 3180
 15 caccgatacg tcaaagatac agagaaacct taatccctcc catagcctag aagcatcaaa 3240
 aagatgattg agcaactcct ccagtaactgg tacattgcac tccctgtatg gttcattctc 3300
 20 cgctacgtgg ctccccacgc acgaaccatc tacttgcgcc acaagctcgg cggggcgccg 3360
 ttcacgcaca cccagtacga cggatggtat ggggtcaagt ttgggcggga gtttctcaa 3420
 25 gcgaagaaga ttggaaggca gacggacttg gtgcacgcgc ggttccgtgg agggggcatg 3480
 gatactttct cgagctatac ttccggcatc catatcattc ttactcggga cccggagaac 3540
 atcaaggcgg tcttgggcgc gaagttegat gaattttcg 3579

30

<210> 17

<211> 31

<212> ADN

35

<213> Secuencia Artificial

<220>

40

<223> cebador

<400> 17

45

ccttaattaa gaggtcgttg gttgagttt c

31

<210> 18

<211> 29

<212> ADN

50

<213> Secuencia Artificial

<220>

55

<223> cebador

<400> 18

60

ccttaattaa ttgataatga cggtgcggg

29

<210> 19

<211> 33

<212> ADN

65

<213> Secuencia Artificial

ES 2 298 276 T3

	<220>		
	<223> cebador		
5	<400> 19		
	aggcgcgccg gagtccaaaa agaccaacct ctg		33
10	<210> 20		
	<211> 34		
	<212> ADN		
	<213> Secuencia Artificial		
15	<220>		
	<223> cebador		
20	<400> 20		
	ccttaattaa tacgtggata ccttcaagca agtg		34
25	<210> 21		
	<211> 35		
	<212> ADN		
	<213> Secuencia Artificial		
30	<220>		
	<223> cebador		
35	<400> 21		
	ccttaattaa gctcacgagt ttgggattt tcgag		35
40	<210> 22		
	<211> 35		
	<212> ADN		
	<213> Secuencia Artificial		
45	<220>		
	<223> cebador		
50	<400> 22		
	gggtttaaac cgcacaggtt ggtcttttg gactc		35
55	<210> 23		
	<211> 10		
	<212> ADN		
	<213> Secuencia Artificial		
60	<220>		
	<223> cebador		
65	<400> 23		
	gggtttaaac		10

ES 2 298 276 T3

	<210> 24	
	<211> 9	
	<212> ADN	
5	<213> Secuencia Artificial	
	<220>	
	<223> cebador	
10	<400> 24	
	aggcgcgcc	9
15	<210> 25	
	<211> 10	
	<212> ADN	
20	<213> Secuencia Artificial	
	<220>	
	<223> cebador	
25	<400> 25	
	ccttaattaa	10
30	<210> 26	
	<211> 21	
	<212> ADN	
35	<213> Secuencia Artificial	
	<220>	
	<223> cebador	
40	<400> 26	
	tcycaaacwg gtacwgcwga a	21
45	<210> 27	
	<211> 21	
	<212> ADN	
50	<213> Secuencia Artificial	
	<220>	
	<223> cebador	
55	<400> 27	
	ggtttgggta aytcwactta t	21
60	<210> 28	
	<211> 20	
	<212> ADN	
65	<213> Secuencia Artificial	

ES 2 298 276 T3

	<220>	
	<223> cebador	
5	<400> 28	
	cggtattayt cyatttcttc	20
10	<210> 29	
	<211> 20	
	<212> ADN	
	<213> Secuencia Artificial	
15	<220>	
	<223> cebador	
20	<400> 29	
	gemcaccrg tacctggacc	20
25	<210> 30	
	<211> 18	
	<212> ADN	
	<213> Secuencia Artificial	
30	<220>	
	<221> cebador	
35	<400> 30	
	atcccaatcg taatcagc	18
40	<210> 31	
	<211> 18	
	<212> ADN	
	<213> Secuencia Artificial	
45	<220>	
	<223> cebador	
50	<400> 31	
	acttgtcttc gtttagca	18
55	<210> 32	
	<211> 18	
	<212> ADN	
	<213> Secuencia Artificial	
60	<220>	
	<223> cebador	
65	<400> 32	
	ctacgtctgt ggtgatgc	18

ES 2 298 276 T3

	<210> 33	
	<211> 25	
	<212> ADN	
5	<213> Secuencia Artificial	
	<220>	
	<223> cebador	
10	<400> 33	
	ctcgggaagc gcgccattgt gttgg	25
15	<210> 34	
	<211> 29	
	<212> ADN	
20	<213> Secuencia Artificial	
	<220>	
	<223> cebador	
25	<400> 34	
	taatacgact cactataggg cgaattggc	29
30	<210> 35	
	<211> 30	
	<212> ADN	
35	<213> Secuencia Artificial	
	<220>	
	<223> cebador	
40	<400> 35	
	gggttaatta acaatactca agcagtttgg	30
45	<210> 36	
	<211> 30	
	<212> ADN	
50	<213> Secuencia Artificial	
	<220>	
	<223> cebador	
55	<400> 36	
	cccttaatta aggggggatg gaagtggccg	30
60	<210> 37	
	<211> 39	
	<212> ADN	
65	<213> Secuencia Artificial	

ES 2 298 276 T3

	<220>		
	<223> cebador		
5	<400> 37		
	ataagaatgc ggccgctgaa cgagaaccac atccaggag		39
10	<210> 38		
	<211> 33		
	<212> ADN		
	<213> Secuencia Artificial		
15	<220>		
	<223> cebador		
20	<400> 38		
	ccttaattaa ggataaccac atccatacgt cgc		33
25	<210> 39		
	<211> 23		
	<212> ADN		
	<213> Secuenciar Artificial		
30	<220>		
	<223> cebador		
35	<400> 39		
	cacaccaccc acgacgactt gtg		23
40	<210> 40		
	<211> 21		
	<212> ADN		
	<213> Secuenciar Artificial		
45	<220>		
	<223> cebador		
50	<400> 40		
	cttcgctgct gaacgactgc g		21
55	<210> 41		
	<211> 46		
	<212> ADN		
	<213> Secuencia Artificial		
60	<220>		
	<221> cebador		
65	<400> 41		
	taatacgact cactataggg aggcacacca cccacgaega cttgtg		46

ES 2 298 276 T3

<210> 42
 <211> 42
 <212> ADN
 5 <213> Secuencia Artificial

 <220>
 <221> cebador
 10
 <400> 42

 ctccgtgct gaacgactgc gaattcttagc gcccttcaag tt 42
 15
 <210> 43
 <211> 222
 <212> ADN
 20 <213> Secuencia Artificial

 <220>
 <221> cebador
 25
 <400> 43

 caggaaacag ctatgaccat gattacgcc aacttggtac cgagctcgga tccactagta 60
 30 acggcccgcca gtgtgctgga attcgccctt aaggccgaat tctgcagata tccatcacac 120
 tggcggccgc tccagcatgc atctagaggg cccaattcgc cctatagtga gtcgtattac 180
 35 aattcactgg ccgtcgtttt acaacgtcgt gactgggaaa ac 222

 <210> 44
 40 <211> 222
 <212> ADN
 <213> Secuencia Artificial

 <220>
 <221> cebador

 <400> 44
 50
 gtcctttgtc gatacggga ctaatgcggg tcgaaccatg gctcgagcct aggtcaccat 60
 tgcggggggg cacacgaact taagccggaa ttcccgotta agacgtctat aggtagctgtg 120
 55 accgcccggc agctcgtagc tagatctccc ggggttaagcg ggatataact cagcataatg 180
 ttaagtgacc ggcagcaaaa tgttgacgca ctgacccttt tg 222
 60
 <210> 45
 <211> 1712
 <212> ADN
 65 <213> *Candida tropicalis*

ES 2 298 276 T3

<400> 45

	ggtaccgagc tcacgagttt tgggattttc gagtttggat tgtttccctt gttgattgaa	60
5	ttgacgaaac cagaggtttt caagacagat aagattgggt ttatcaaac gcagttcgaa	120
	atattccagt tgggtttccaa gatattcttga agaagattga cgatttgaaa ttgagagag	180
10	tggagaagat ctgggttggg ttgttggaga atttcaagaa tctcaagatt tacttcaacg	240
	acgggtacaa cgagaattgt attgaattga tcaagaacat gatcttgggt ttacagaaca	300
	tcaagttctt ggaccagact gagaatgcc aagatataca aggcgtcatg tgataaatg	360
15	gatgagattt atcccacaat tgaagaaaga gtttatggaa agtgggtcaac cagaagctaa	420
	acaggaagaa gcaaacgaag aggtgaaaca agaagaagaa ggtaaataag tattttgcat	480
20	tatataacaa acaaagtaag gaatacagat ttatacaata aattgccata ctagtcacgt	540
	gagatatctc atccattccc caactcccaa gaaaaaaaaa aagtgaaaaa aaaaatcaaa	600
25	cccaaagatc aacctcccca tcatcatcgt catcaaacc ccagctcaat tcgcaatggg	660
	tagcacaaaa acatacacag aaagggcatc agcacacccc tccaagggtg cccaacgttt	720
30	attccgctta atggagtcca aaaagaccaa cctctgcgcc tcgatcgacg tgaccacaac	780
	cgccgagttc ctttcgctca tcgacaagct cgggtccccc atctgtctcg tgaagacgca	840
35	catcgatata atctcagact tcagctacga gggcacgatt gagccgttgc ttgtgcttgc	900
	agagcgccac gggttcttga tattcgagga caggaagttt gctgatatcg gaaacaccgt	960
40	gatgttgacg tacacctcgg gggatatacc gatcgccggc tggagtgaac tcacgaacgc	1020
	gcacggagtg actgggaagg gcgtcgttga aggggtgaaa cgcggtgcgg agggggtaga	1080
	aaaggaaagg ggcgtgttga tgttgccgga gttgtcagat aaaggctcgt tggcgcattg	1140
45	tgaatatacc cgtgagacga tcgagattgc gaagagtgat cgggagttcg tgattgggtt	1200
	catcgcgacg cgggacattg ggggtagaga agaagggttt gattggatca tcatgacgcc	1260
50	tgggtgtggg ttggatgata aaggcgatgc gttgggccag cagtatagga ctgttgatga	1320
	ggtggttctg actggtaccg atgtgattat tgcgggaga gggttgtttg gaaaaggaag	1380
55	agaccctgag gtggagggaa agagatacag gcatgctgga tggaaagcat acttgaagag	1440
	aactggtcag ttagaataaa tattgtaata aataggtcta tatacataca ctaagcttct	1500
	aggacgtcat tgtagtcttc gaagttgtct gctagtttag ttctcatgat ttcgaaaacc	1560
60	aataacgcaa tggatgtagc agggatggtg gttagtgcgt tctgacaaa ccagaggtac	1620
	gccgcctcaa accacgtcac attcgccctt tgcttcatcc gcataccttg cttgaaggta	1680
65	tccacgtacg agttgtaata caccttgaag aa	1712