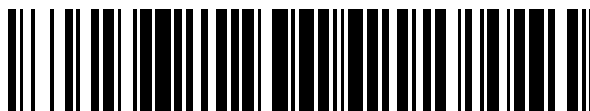


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 379 776**

51 Int. Cl.:  
**C12N 9/00** (2006.01) **C07K 1/00** (2006.01)  
**C12N 1/20** (2006.01)  
**C12N 15/00** (2006.01)  
**C12Q 1/00** (2006.01)  
**C12Q 1/34** (2006.01)  
**C12Q 1/68** (2006.01)  
**C12P 21/06** (2006.01)  
**C12P 21/04** (2006.01)  
**A61K 38/44** (2006.01)  
**C07H 21/04** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **03724119 .7**  
96 Fecha de presentación: **18.04.2003**  
97 Número de publicación de la solicitud: **1576095**  
97 Fecha de publicación de la solicitud: **21.09.2005**

54 Título: **Genes y proteínas de oxidasa de alcohol graso de *Cándida tropicalis* y métodos relacionados con los mismos**

30 Prioridad:  
**19.04.2002 US 374021 P**  
**18.04.2003 US 418819**

45 Fecha de publicación de la mención BOPI:  
**03.05.2012**

45 Fecha de la publicación del folleto de la patente:  
**03.05.2012**

73 Titular/es:  
**COGNIS IP MANAGEMENT GMBH**  
**HENKELSTRASSE 67**  
**40589 DÜSSELDORF, DE**

72 Inventor/es:  
**EIRICH, L. Dudley y**  
**CRAFT, David, L.**

74 Agente/Representante:  
**Carvajal y Urquijo, Isabel**

ES 2 379 776 T3

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

## DESCRIPCIÓN

Genes y proteínas de oxidasa de alcohol graso de *Candida tropicalis* y métodos relacionados con los mismos

Referencia cruzada con solicitudes relacionadas

5 La presente invención reivindica el beneficio de la Solicitud Provisional de los Estados Unidos con Serial No. 60/374.021, presentada el 19 de abril de 2002, la cual se incorpora aquí como referencia.

Antecedentes de la invención

10 Diferentes cepas modificadas por ingeniería genética de *Candida tropicalis* han sido utilizadas en fermentaciones para la bioconversión del ácido graso de 18 carbonos como ácido oleico en ácido dicarboxílico de 18 carbonos. Cuando se cultiva sobre ácidos grasos, *C. tropicalis* de tipo silvestre convierte ácidos grasos de cadena larga en acetil CoA por medio de un proceso conocido como oxidación en  $\beta$ , que es el catabolismo secuencial de fragmentos de 2 carbonos de longitud de un ácido graso en acetil CoA. La oxidación en  $\beta$  es por lo tanto llamada así debido a que se presenta el ataque oxidativo inicial en el segundo átomo de carbono del grupo carboxílico. *C. tropicalis* puede catabolizar también ácidos grasos a través de una ruta de oxidación en  $\omega$  en la cual únicamente el carbono del metilo terminal es oxidado hasta un ácido carboxílico, produciendo un ácido dicarboxílico. En la oxidación en  $\omega$ , el ácido graso es convertido en ácido dicarboxílico a lo largo de una ruta de tres etapas comenzando con la oxidación del grupo metilo terminal hasta un alcohol. Esta etapa es catalizada por el complejo hidroxilasa que contiene tanto las proteínas citocromo P450 monooxigenasa (CYP) como la citocromo P450 reductasa (NCP). El alcohol es luego convertido en un aldehído por medio de la alcohol graso oxidasa (FAO) luego en el ácido dicarboxílico por medio de una aldehído deshidrogenasa. El producto deseado es el ácido dicarboxílico de cadena larga. Una oxidasa de alcohol graso se diferencia de una alcohol oxidasa en su especificidad por una cadena larga. Las alcohol oxidasas son en general específicas para metanol pero algunas veces pueden oxidar alcoholes hasta de C4. Las oxidasas de alcohol graso generalmente no oxidan alcoholes con longitudes de cadena menores de ocho.

25 En la *Candida tropicalis* de tipo silvestre, la oxidación en  $\beta$  consume ácidos grasos mucho más rápido de lo que la ruta de la oxidación en  $\omega$  puede oxidarlos. Sin embargo, por medio de la inactivación de los genes *POX 4* y *POX 5*, cuyos productos génicos son responsables por la iniciación de la oxidación en  $\beta$ , tales cepas modificadas por medio de ingeniería genética de *C. tropicalis* preferencialmente desvían los ácidos grasos dentro de la ruta de la oxidación en  $\omega$ . La cepa base utilizada para el desarrollo de diversas cepas amplificadas por el gen es H5343, que es la cepa 20336 de *C. tropicalis* (American Type Culture Collection) tanto con los genes *POX 4* como *POX 5* inactivados por medio de inactivación por inserción. La cepa primaria utilizada en fermentaciones de producción a gran escala es HDC23-3, que se deriva de H5343, pero también tiene al gen *CYP52A2* (una citocromo P450 monooxigenasa) amplificado. El complejo de hidroxilasa es responsable por catalizar la primera etapa en la oxidación en  $\omega$ , que es considerada la etapa limitante de la velocidad. La amplificación de los genes *CYP52A2* y *NCP* ayuda a superar esta limitación de la velocidad, pero luego el siguiente cuello de botella vuelve la conversión del alcohol en el aldehído por medio de la enzima FAO. Durante las fermentaciones con HDC23-3, se ha descubierto que una pequeña cantidad (aproximadamente 0,5% p/p en el caldo) del ácido graso  $\omega$ -hidroxilado (HFA) se acumula durante la fermentación. Este producto de oxidación parcial interfiere con las etapas de purificación posteriores y provoca rendimientos totales menores. Existe por lo tanto la necesidad adicional, de reducir el cuello de botella en la conversión del alcohol en un aldehído durante la segunda etapa de la ruta de oxidación en  $\omega$ .

40 Un pequeño número de oxidasas de alcohol graso ha sido descrito en la literatura científica en diversas levaduras, ejemplos de las cuales son *Candida tropicalis* (1,2,3,4), *Candida maltosa* (5,6), *Candida cloacae* (4), *Torulopsis candida* (7), *Candida (Torulopsis) bombicola* (8), y *Candida (Torulopsis) apicola* (9). La FAO fue purificada a partir de la levadura cultivada en hexadecano, *T. candida* (7) y descrita como un tetrámero (pm 290 kD) con un peso molecular por subunidad de 75 ID. Tiene un pH óptimo de 7,6 y oxida alcoholes superiores con una cadena carbonada de C4 a C16. *C. bombicola* (8) cultivada en hexadecano tiene aparentemente dos diferentes actividades como alcohol oxidasa, una con una especificidad óptima de longitud de cadena de 10 para n-alcoholes y otra con una especificidad óptima de longitud de cadena de 14. La FAO de *C. maltosa* (6) cataliza la oxidación de 1-alcoholes (C4 a C22) con la actividad más alta utilizando 1-octanol. También oxida 2-alcoholes (C8 a C16). Se encontró que los  $\alpha,\omega$ -alcanodiolos, el ácido  $\omega$ -hidroxipalmítico, los fenilalcoholes y los alcoholes terpénicos, todos eran sustratos para la FAO, pero con muy bajas tasas de oxidación. La oxidación de 2-alcoholes es estereoselectiva únicamente para los R(-) enantiómeros.

55 La FAO de *C. tropicalis* (ATCC 20336) cultivada sobre hexadecano fue descrita primero por Kemp et al. en 1988 (1). Se encontró que oxida 1-alcoholes de C4 a C18, pero tiene una actividad máxima con dodecanol. Se encontró que oxida 16-hidroxipalmitato pero no 12-hidroxilaurato. La FAO fue posteriormente purificada (3) y se demostró que era un dímero (pm = 145 kD) con un peso molecular por subunidad de 68 - 72 kD. La enzima purificada mostró especificidad similar por el sustrato como se describió anteriormente, pero demostró actividad adicional con 12-hidroxilaurato y 2-dodecanol. Se encontró que la enzima era una flavoproteína sensible a la luz, pero la identidad de

la flavina no se conocía (10).

5 Recientemente se clonaron dos genes *FAO* de *C. cloacae* y un gen *FAO* de *C. tropicalis* y se determinaron las secuencias de ADN (4, 12). Las identidades del marco de lectura abierto (ORF) y 74,8% y 76,2% similitudes de la secuencia de aminoácidos con la *FAO1* y *FAO2* de *C. cloacae*, respectivamente. El gen *FAO1* pero no el gen *FAO2* ha sido exitosamente clonado y expresado en *Escherichia coli*.

WO 99/47685 divulga secuencias aisladas de ácido nucleico que codifican un polipéptido que tiene actividad de *FAO*, constructos de supresión de ácido nucleico, células mutantes que suprimen *FAO*, y usos de los mismos.

10 La presente invención proporciona genes *FAO* de *C. tropicalis* y composiciones y métodos que emplean los genes *FAO*. Las composiciones y los métodos son útiles para incrementar la actividad de la *FAO* durante la segunda etapa de oxidación en omega de ácido grasos y en última instancia en un incremento en productividad de diácido.

Breve descripción de los dibujos

La Figura 1 es un mapa de restricción de *FAO1jf* que es la región de codificación del gen para *FAO1* ligado dentro del vector de expresión pJF118EH.

15 La Figura 2 es un mapa de restricción de *FAO2jf* que es la región de codificación del gen para *FAO2* ligado dentro del vector de expresión pJF118EH.

La Figura 3 describe gráficamente la actividad típica de la alcohol oxidasa en fermentaciones a escala de laboratorio con ácido oleico como sustrato.

La Figura 4 describe gráficamente la actividad de *FAO1*, *FAO2a*, y *FAO2a'* sobre 1-alcanoles.

La Figura 5 describe gráficamente la actividad de *FAO 1*, *FAO2a* y *FAO2a'* sobre 2-alcanoles.

20 La Figura 6 describe gráficamente la actividad de *FAO1*, *FAO2*, y *FAO2a'* sobre otros alcanoles.

La Figura 7 describe gráficamente una comparación de productividad en fermentaciones con cepas amplificadas por la *FAO* o una cepa base, H5343.

La Figura 8 describe gráficamente una comparación de la concentración del ácido graso T-hidroxi en fermentaciones con la cepa amplificada por la *FAO*, HDC40-7, o la cepa base, H5343.

25 La Figura 9 describe gráficamente una comparación de la actividad de la alcohol oxidasa entre la cepa amplificada por *FAO1*, HDC40-7, y la cepa base, H5343.

La Figura 10 describe gráficamente una comparación de la concentración del ácido graso T-hidroxi en fermentaciones con la cepa amplificada por la *FAO*, HDC40-1, HDC40-5, o HDC40-7, utilizando ácido ricinoleico como sustrato.

30 Resumen de la invención

35 La presente invención proporciona una oxidasa de alcohol graso o un fragmento enzimáticamente activo de la misma que tiene una identidad de secuencia de aminoácidos de al menos una superior a NO: 2, o superior al 95% cuando se la compara con la secuencia de aminoácidos expuesta en la SEQ ID NO: 2. Preferiblemente, esta oxidasa de alcohol graso incluye además un péptido firma que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 13.

De acuerdo con la presente invención, se proporciona también una molécula aislada de ácido nucleico que contiene un marco de lectura abierto (ORF) para un gen para la oxidasa de alcohol graso (*FAO*) de *Candida tropicalis* en donde el ORF está operativamente enlazado a un promotor que es capaz de afectar la expresión del ORF. Preferiblemente, la *FAO* incluye la secuencia de aminoácidos expuesta en la SEQ ID NO: 2.

40 Se proporcionan también vectores que contienen las moléculas aisladas de ácido nucleico descritas aquí. Tales vectores incluyen pero no se limitan a vectores plásmidos, fagémidos, fagos, cósmidos, o de ADN lineal. Además, también se proporcionan células huésped las cuales contienen un vector objetivo o una molécula aislada de ácido nucleico. Los ejemplos de células huésped incluyen células bacterianas, células de hongos, células de insecto, células de animales o células de plantas. Los ejemplos de células de hongos incluyen a aquellas de *Yarrowia sp.*,

*Bebaromyces sp.*, *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.* y *Candida sp.* Se prefieren especialmente las células huésped de *Candida tropicalis*.

#### Descripción detallada de la invención

5 La productividad del diácido se mejora de acuerdo con la presente invención incrementando selectivamente las enzimas que se sabe que son importantes para la oxidación de sustratos orgánicos tales como ácidos grasos o compuestos alifáticos, por ejemplo alcanos. De acuerdo con la presente invención, se han identificado y caracterizado dos genes para oxidasa de alcohol graso (FAO) de *C. tropicalis*. Los genes FAO codifican las enzimas oxidasa de alcohol graso, que catalizan la conversión de un alcohol en un aldehído durante la segunda etapa de la ruta de oxidación en  $\omega$ . La amplificación del número de copias del gen FAO y/o la actividad transcripcional en una  
10 célula huésped da como resultado una mayor actividad de la alcohol oxidasa, una mayor productividad de diácido, y, cuando se utilizan sustratos de ácido graso, menores niveles de ácido graso  $\omega$ -hidroxi.

15 La presente invención provee secuencias de nucleótidos de *C. tropicalis* que codifican dos diferentes oxidasas de alcohol graso, cada una de las cuales tiene una especificidad diferente del sustrato. En una modalidad, se provee una molécula aislada de ácido nucleico que codifica a la enzima oxidasa de alcohol graso FAO1, que tiene una secuencia de aminoácidos como la expuesta en la SEQ ID NO: 2. Un ejemplo de tal molécula aislada de ácido nucleico incluye al gen FAO1 que tiene la secuencia de nucleótidos como la expuesta en la SEQ ID NO: 1.

20 En otro aspecto de la invención, se proveen moléculas aisladas de ácido nucleico que codifican a la enzima oxidasa de alcohol graso FAO2. De acuerdo con la presente invención, se han identificado y aislado dos alelos que codifican a la enzima FAO2. Los dos alelos, FAO2a y FAO2b, son 95% idénticos por la secuencia de ADN y tienen 98% de similitud por la secuencia de aminoácidos. La enzima FAO2a incluye la secuencia de aminoácidos como la expuesta en la SEQ ID NO: 4. Un ejemplo de una secuencia de nucleótidos que codifica la enzima FAO2a es expuesta en la SEQ ID NO: 3. La enzima FAO2b incluye la secuencia de aminoácidos como la expuesta en la SEQ ID NO: 6. Un ejemplo de una secuencia de nucleótidos que codifica la enzima FAO2b es expuesta en la SEQ ID NO: 5.

25 Recientemente se ha determinado que ciertos eucariotas, por ejemplo, ciertas levaduras, no se adhieren, en algunos aspectos, al código genético "universal" que establece que codones particulares (tripletes de ácidos nucleicos) codifican para aminoácidos específicos. En realidad, el código genético es "universal" debido a que es virtualmente el mismo en todos los organismos vivos. Se sabe que ciertos *Candida sp.* traducen el codón CTG (que, de acuerdo con el código "universal" designa leucina), como serina. Ver, por ejemplo, Ueda et al., Biochemie (1994) 76, 1217 - 1222, donde se demuestra que *C. tropicalis*, *C. cylindracea*, *C. guilliermodii* y *C. lusitanae* se adhieren al  
30 código "no universal" con respecto al codón CTG. Por lo tanto, las secuencias de ácido nucleico pueden codificar para una secuencia de aminoácidos en organismos del código "universal" y una variante de esa secuencia de aminoácidos en organismos de código "no universal" dependiendo de la presencia de codones CTG en la secuencia de codificación de ácido nucleico. La diferencia puede llegar a ser evidente cuando, en el transcurso de la modificación por ingeniería genética, una molécula de ácido nucleico que codifica una proteína es transferida de un  
35 organismo de código "no universal" a un organismo de código "universal".

Por lo tanto, la presente invención también provee una secuencia de aminoácidos (expuesta en la SEQ ID NO: 10) para una enzima FAO2a cuando FAO2a es expresada en una especie de *Candida* por ejemplo *C. tropicalis*.

40 La presente invención también provee proteínas FAO. Por ejemplo, se provee una proteína FAO1 que tiene una secuencia de aminoácidos como la expuesta en la SEQ ID NO: 2 o un fragmento enzimáticamente activo de la misma que tiene una identidad de secuencia de aminoácidos al menos superior al 90% cuando se compara con la secuencia de aminoácidos expuesta en la SEQ ID NO : 2. Como se utiliza aquí, "fragmento enzimáticamente activo" se refiere a una porción de la secuencia de aminoácidos de FAO como la expuesta en la SEQ ID NO: 12 o se provee un fragmento enzimáticamente activo de la misma. Como se utiliza aquí, "fragmento enzimáticamente activo" se refiere a una porción de la enzima FAO que es suficiente para retener actividad enzimática en la conversión de un  
45 alcohol en un aldehído.

De acuerdo con la presente invención, las proteínas FAO1 y FAO2 pueden ser preparadas por medio de métodos familiares para aquellos capacitados en el arte tales como por medio de clonación de los genes FAO1 y FAO2 (incluidos los alelos FAO2a y/o FAO2b) en un vector de expresión apropiado seguido por expresión en una célula huésped adecuada. Ver, por ejemplo, la patente de los Estados Unidos No. 6.331.420, incorporada aquí como referencia como si estuviera expuesta en su totalidad. La enzima relevante o fragmento de la misma, puede ser también generada por medio de amplificación directa de la correspondiente secuencia de codificación a través de PCR, seguido por procedimientos recombinantes estándar y expresión en una célula huésped adecuada. Con respecto a análogos de FAO, derivados, moléculas tipo FAO, porciones o fragmentos enzimáticamente activos de la misma (como se define aquí), se pueden diseñar iniciadores de PCR para permitir amplificación directa de  
50 secuencias de codificación para la correspondiente inserción, supresión o sustitución de aminoácidos de las variantes de aminoácidos (como se define aquí). Los iniciadores para uso en PCR pueden ser oligonucleótidos

sintéticos preparados en un sintetizador automatizado de oligonucleótidos tal como un sintetizador de ADN ABI de Perkin-Elmer Corporation. Además, los oligonucleótidos pueden ser adquiridos a fabricantes comerciales, por ejemplo, a Synthetic Genetics (San Diego, Calif.). Las proteínas FAO1 y FAO2 pueden también ser sintetizadas químicamente utilizando metodologías bien conocidas.

5 La secuencia apropiada de ADN puede ser insertada en un vector por medio de una variedad de procedimientos que se considera que están dentro del alcance de aquellos capacitados en el arte, que pueden incluir inserción del ADN en un(os) sitio(s) apropiado(s) de endonucleasa de restricción o clonación de la secuencia de ADN en los vectores de expresión utilizando reacción en cadena de la polimerasa de alta fidelidad. Las técnicas estándar para la construcción de tales vectores son bien conocidas por aquellos ordinariamente capacitados en el arte y pueden encontrarse en referencias tales como Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, o cualquiera entre la gran cantidad de manuales de laboratorio sobre tecnología de ADN recombinante que se encuentran fácilmente disponibles. Una variedad de estrategias se encuentran disponibles para ligar fragmentos de ADN, cuya escogencia dependen de la naturaleza de los terminales de los fragmentos de ADN. El vector puede incluir también secuencias apropiadas para la expresión de la amplificación, o secuencias que facilitan la clonación, expresión o procesamiento.

Además, los vectores de expresión contienen preferiblemente uno o más genes marcadores seleccionables para proporcionar un rasgo fenotípico para selección de células huésped transformadas tales como resistencia a la dihidrofolato reductasa o neomicina para cultivo de células eucariotas, o de resistencia a la tetraciclina o a la ampicilina en *E. coli*.

20 El vector que contiene la secuencia apropiada de ADN como se describió aquí anteriormente, así como un promotor apropiado o secuencia de control, puede ser empleado para transformar un huésped apropiado para permitir que el huésped exprese la enzima. "Transformación" incluye todas las formas que causan la incorporación de ADN foráneo por una célula huésped. Las células transformadas son luego seleccionadas por aquellas que contienen el ADN deseado y se cultivan los transformantes exitosos bajo condiciones que afecten la expresión de las secuencias de codificación.

Los ejemplos representativos de huéspedes apropiados incluyen células bacterianas, tales como *E. coli* y *Streptomyces*; células de hongos, tales como levadura; células de insecto tales como *Drosophila* S2 y *Spodoptera* Sf9; células de animales; adenovirus; células de plantas, etc. Se considera que la selección de un huésped apropiado se encuentra dentro del alcance de aquellos capacitados en el arte a partir de las enseñanzas expuestas aquí y pueden incluir a la cepa huésped para expresión BL21 o BL21 CodonPlus RIL (Stratagene, La Jolla, CA).

Después de la transformación de la cepa huésped, se induce la enzima por los medios apropiados (por ejemplo, cambio de temperatura o inducción química) y se cultivan las células durante un periodo de tiempo adicional. Si la secuencia que codifica a la FAO está bajo el control de su promotor nativo, se puede utilizar ácido oleico para inducir expresión. Si la FAO está operativamente enlazada a un promotor heterólogo, se pueden utilizar otros inductores. Un método preferido de inducción es a través del uso de IPTG (Isopropil-β-D-tiogalactopiranosido).

Típicamente se rompen las células por medios físicos o químicos conocidos por aquellos capacitados en el arte, se centrifugan, y se retiene el extracto crudo resultante para purificación adicional.

La purificación de FAO1 o FAO2 o de un fragmento de las mismas se puede llevar a cabo a través de cualquier medio conocido por aquellos capacitados en el arte y puede incluir técnicas cromatográficas, tales como cromatografía de afinidad con etiqueta de histidina y cromatografía de afinidad con níquel. Los kits comercialmente disponibles para esta purificación pueden ser adquiridos a los proveedores (Qiagen, Inc. Chatsworth, CA; Novagen, CA, EUA). Se puede utilizar luego un inmunoensayo, tal como un ensayo de transferencias tipo Western, para verificar la presencia de la enzima recombinante.

También de acuerdo con la presente invención, se han descubierto tres diferentes secuencias de péptidos firma que son únicas para las proteínas FAO1 y FAO2 de la presente invención. El primer péptido firma tiene la secuencia de aminoácidos: CGFCYLGC (SEQ ID NO: 13). Este primer péptido firma se encuentra tanto en proteínas FAO1 como FAO2 de la presente invención pero no en otras proteínas FAO de levadura previamente caracterizadas. Con referencia a la SEQ ID NO: 2 (proteína FAO1), la SEQ ID NO: 4 (proteína FAO2a), y las posiciones de los aminoácidos 355 a 362 de la SEQ. Un segundo péptido firma tiene la secuencia de aminoácidos IIGSGAGAGVMA (SEQ ID NO: 14). Este segundo péptido firma está presente en las proteínas FOA2a y FAO2b de la presente invención pero no en la proteína FAO1 de la presente invención, ni en las proteínas FAO de levadura, previamente caracterizadas. Con referencia a las SEQ ID Nos: 4 y 6, el segundo péptido firma se localiza en las posiciones de los aminoácidos 198 a 209. Un tercer péptido firma se encuentra en las proteínas FAO2a y FAO2b de la presente invención pero no en otras proteínas FAO de levadura previamente caracterizadas. El tercer péptido firma tiene la secuencia de aminoácidos AGSTLGGG (SEQ ID NO: 19). Con referencia a las SEQ ID Nos: 4 y 6, este tercer péptido firma se localiza en las posiciones de los aminoácidos 262 a 269.

Por lo tanto, de acuerdo con la presente invención, se provee un péptido que tiene la secuencia de aminoácidos: CGFCYLGC (SEQ ID NO: 13). Estos péptidos son útiles, por ejemplo, en la producción de anticuerpos que se enlazan específicamente con una FAO de levadura que contiene la secuencia de aminoácidos de los péptidos firma. Tales anticuerpos son útiles en inmunoensayos por ejemplo radioinmunoensayos, ensayos por inmunoadsorción ligados a enzimas (ELISA), transferencias tipo Western, ensayos de inmunofluorescencia, ensayos de quimioluminiscencia y ensayos de bioluminiscencia. Estos tipos de ensayos son útiles para monitorear los niveles de enzima FAO en diferentes momentos durante un procedimiento de fermentación y pueden ayudar a resolver problemas que puedan surgir durante la fermentación.

Las secuencias de aminoácidos estructuralmente relacionadas pueden ser sustituidas por las secuencias divulgadas expuestas en las SEQ ID Nos: 2, 4, 6, 10, 12, 13, 14, ó 19 en la realización de la presente invención. Los derivados por inserción de aminoácidos de las proteínas y péptidos de la presente invención incluyen fusiones en el terminal amino y/o carboxilo así como inserciones dentro de la secuencia de un solo o de múltiples aminoácidos. Las variantes por inserción en la secuencia de aminoácidos son aquellas en las cuales se introducen uno o más residuos de aminoácidos en un sitio predeterminado en una proteína o péptido FAO objetivo aunque también es posible una inserción al azar con una selección adecuada del producto resultante. Las variantes por supresión se pueden elaborar removiendo uno o más aminoácidos de la secuencia de un péptido objetivo. Las variantes por sustitución de aminoácidos son aquellas en las cuales se ha removido al menos un residuo en la secuencia y se ha insertado en su lugar un residuo diferente. Las sustituciones típicas son aquellas elaboradas de acuerdo con la Tabla 1 siguiente:

Tabla 1

Residuos adecuados para sustituciones de aminoácidos	
Residuo Original	Ejemplos de Sustituciones
Ala (A)	Ser
Arg (R)	Lys
Asn (N)	Gln; His
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
Gly (G)	Pro
His (H)	Asn; Gln
Ile (I)	Leu; Val
Leu (L)	Ile; Val
Lys (K)	Arg; Gln; Glu
Met (M)	Leu; Ile
Phe (F)	Met; Leu; Tyr
Ser (S)	Thr

(continuación)

Residuos adecuados para sustituciones de aminoácidos	
Residuo Original	Ejemplos de Sustituciones
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp; Phe
Val (V)	Ile; Leu

5 Cuando una proteína o péptido FAO objetivo se obtiene por derivación a partir de sustitución de aminoácidos, los aminoácidos son generalmente reemplazados por otros aminoácidos que tienen propiedades tales como hidrofobicidad, hidrofiliidad, electronegatividad, cadenas laterales voluminosas y similares. Como se utiliza aquí, los términos "derivado", "análogo", "fragmento", "porción" y "moléculas similares" se refieren a una proteína FAO objetivo que tiene una secuencia de aminoácidos como la expuesta en la SEQ ID No: 2 y que tiene una sustitución, inserción, adición, o supresión de aminoácidos con tal de que dicho derivado, análogo, fragmento, porción, o molécula similar retenga la habilidad para actuar como una oxidasa de alcohol graso, es decir, que retenga la capacidad para convertir un alcohol en un aldehído. Igualmente, los términos "derivado", "análogo", "fragmento", "porción" y "moléculas similares" pueden referirse también a un péptido firma FAO objetivo que tiene una secuencia de aminoácidos como la expuesta en las SEQ ID Nos: 13, 14, ó 19 y que tiene una sustitución, inserción, adición, o supresión de aminoácidos, con tal de que dicho derivado, análogo, fragmento, porción, o molécula similar de un péptido firma FAO, cuando se incorpora dentro de una proteína FAO mayor, derivado, análogo, fragmento, porción, o molécula similar, no disminuya la actividad de FAO.

Los péptidos sintéticos de la presente invención pueden ser sintetizados por medio de una variedad de técnicas conocidas. Por ejemplo, los péptidos pueden ser preparados utilizando la técnica en fase sólida inicialmente descrita por Merrifield (1963) in J. Am. Chem. Soc. 85:2149 - 2154. Otras técnicas de síntesis de péptidos pueden ser encontradas en M. Bodanszky et al. Peptide Synthesis, John Wiley and Sons, 2d Ed., (1976) y otras referencias fácilmente disponibles para aquellos capacitados en el arte. Un resumen de las técnicas de síntesis de péptidos puede ser encontrado en J. Sturart y J.S. Young, Solid Phase Peptide Synthesis, Pierce Chemical Company, Rockford, Ill., (1984). Los péptidos también pueden ser sintetizados por medio de métodos en solución como se describe en The Proteins, Vol. II, 3d Ed., Neurath, H. et al., Eds., pp. 105 - 237, Academic Press, New York, N.Y. (1976). Los grupos de protección apropiados para uso en diferentes síntesis de péptidos están descritos en los textos enlistados anteriormente así como en J.F.W. McOmie, Protective Groups in Organic Chemistry, Plenum Press, New York, N.Y. (1973). Los péptidos de la presente invención pueden ser preparados también por medio de escisión química o enzimática de porciones mayores de una proteína FAO objetivo o a partir de una proteína FAO de longitud completa.

30 Adicionalmente, las proteínas y péptidos FAO de la presente invención pueden ser preparados también por medio de técnicas de ADN recombinante. Para la mayoría de los aminoácidos utilizados para construir proteínas, más de una tripleta de nucleótidos de codificación (codón) pueden codificar para un residuo aminoácido particular. Esta propiedad del código genético es conocida como redundancia. Por lo tanto, una cantidad de secuencias diferentes de nucleótidos pueden codificar para una proteína o péptido FAO particular.

35 De este modo, otras proteínas FAO o fragmentos enzimáticamente activos de las mismas, que comparten suficiente identidad de aminoácidos con la FAO1, FAO2a, y FAO2b de la presente invención (SEQ ID Nos: 2, 4, y 6) se encuentran dentro del alcance de la presente invención.

40 En una modalidad, una enzima FAO objetivo tiene una identidad de secuencia de aminoácidos superior al 90% cuando se la compara con la secuencia de aminoácidos de FAO1 como la expuesta en la SEQ ID NO: 2. En esta modalidad, es preferible que una enzima FAO objetivo que tiene una identidad de aminoácidos superior al 90% cuando se la compara con la secuencia de aminoácidos de FAO1 como la expuesta en la SEQ ID NO: 2 también incluya el motivo firma GGFCYLGC (SEQ ID NO: 13). En una modalidad más preferida, la enzima FAO tiene una identidad de secuencia superior al 95% cuando se la compara con la secuencia de aminoácidos de FAO1 como la expuesta en la SEQ ID NO: 2. En esta modalidad, es preferible que la enzima FAO objetivo tenga una identidad de aminoácidos superior al 95% cuando se la compara con la secuencia de aminoácidos de FAO1 como la expuesta en la SEQ ID NO: 2 también incluya el motivo firma GGFCYLGC (SEQ ID NO: 13).

La presente invención también provee moléculas de ácido nucleico que contienen secuencias de nucleótidos que codifican para el motivo firma de una FAO1 y FAO2 objetivo. Por ejemplo, una secuencia de nucleótidos que codifica al motivo firma CGFCYLGC (SEQ ID NO: 13) se presenta como:

TGY GGN TTY TGY TAY YTN GGN TGY (SEQ ID NO: 32) en donde:

- 5 R es A o G, Y es C o T, M es A o C, K es G o T, S es G o C, W es A o T, H es A o T o C, B es G o T o C, D es G o A o T, N es A o G, o C o T.

Además, se presenta una molécula de ácido nucleico que contiene una secuencia de nucleótidos que codifica para el motivo firma IIGSGAGAGVMA (SEQ ID NO: 14) de una FAO2 objetivo que tiene la siguiente secuencia:

ATH ATH GGN WSN GGN GCN GGN GCN GGN GTN ATG GCN (SEQ ID NO: 33) en donde:

- 10 R es A o G, Y es C o T, M es A o C, K es G o T, S es G o C, W es A o T, H es A o T o C, B es G o T o C, D es G o A o T, N es A o G, o C o T.

Además, se presenta una molécula de ácido nucleico que contiene una secuencia de nucleótidos que codifica para el motivo firma AGSTLGGG (SEQ ID NO: 19) como:

GCN GGN WSN ACN YTN GGN GGN GGN (SEQ ID NO: 34)

- 15 en donde R es A o G, Y es C o T, M es A o C, K es G o T, S es G o C, W es A o T, H es A o T o C, B es G o T o C, D es G o A o T, N es A o G, o C o T.

Las moléculas de ácido nucleico que codifican a los motivos firma FAO1 y FAO2 son útiles para identificar y aislar genes que codifican proteínas FAO de *Candida tropicalis*.

- 20 Los métodos de alineación de nucleótidos y de secuencias de aminoácidos para comparación son bien conocidos en el arte. De este modo, la determinación del porcentaje de identidad entre dos secuencias cualesquiera se puede lograr utilizando un algoritmo matemático. Las implementaciones de ordenador de estos algoritmos matemáticos pueden ser utilizadas para comparación de las secuencias para determinar la identidad de la secuencia. Tales implementaciones incluyen, pero no se limitan a: CLUSTAL en el programa PC/Gene (de Intelligenetics, Mountain View, California); el programa ALIGN (Versión 2.0) y GAP, BESTFIT, BLAST, FASTA, y TFASTA en el Wisconsin Genetics Software Package, Versión 8 (de Genetics Computer Group (CGC), 575 Science Drive, Madison, Wisconsin, EUA). FASTA, y TFASTA en el Wisconsin Genetics Software Package, Versión 8 (de Genetics Computer Group (CGC), 575 Science Drive, Madison, Wisconsin, EUA). Las alineaciones que utilizan estos programas pueden ser llevadas a cabo utilizando los programas predeterminados. El programa CLUSTAL está bien descrito por Higns et al., 1988, Higgins et al., 1989, Corpet et al. 1958, Huang et al. 1992, y Pearson et al., 1994. El programa para realizar los análisis por BLAST se encuentra públicamente disponible a través del National Center for Biotechnology Information (). Para los propósitos de la presente invención, la comparación de las secuencias de nucleótidos y de aminoácidos para la determinación del porcentaje de identidad de secuencia con las secuencias de aminoácidos y nucleótidos de FAO1 y FAO2, se realiza preferiblemente utilizando el programa BlastN (versión 1.4.7 o posterior) con sus parámetros predeterminados o cualquier programa equivalente. Por "programa equivalente" se entiende cualquier programa de comparación de secuencia que, para dos secuencias cualesquiera en cuestión, genere una alineación que tenga correspondencias idénticas de nucleótidos o de residuos aminoácidos y un porcentaje idéntico de identidad de secuencia cuando se la compara con la correspondiente alineación generada por el programa preferido.
- 25
- 30
- 35

- 40 Otra forma de describir moléculas aisladas de ácido nucleico que codifican una FAO objetivo es en términos de hibridación con las secuencias de codificación (ORF) de los genes FAO1 (SEQ ID NO: 2) y FAO2a y FAO2b objetivo (SEQ ID Nos: 4 y 6, respectivamente). Por lo tanto, de acuerdo con la presente invención, se provee una molécula de ácido nucleico que hibrida bajo condiciones de rigurosidad media y alta con la secuencia de codificación del gen FAO1 objetivo, es decir, los nucleótidos 1941 - 4054 de la secuencia de nucleótidos expuesta en la SEQ ID NO: 1 y que también codifica al motivo firma CGFCYLGC (SEQ ID NO: 13). Preferiblemente, la molécula de ácido nucleico que hibrida bajo rigurosidad media a alta en donde:
- 45

R es A o G, Y es C o T, M es A o C, K es G o T, S es G o C, W es A o T, H es A o T o C, B es G o T o C, D es G o A o T, N es A o G, o C o T.

- 50 Como se utiliza aquí, la hibridación bajo condiciones de rigurosidad media o alta son como se describe en Maniatis et al. 1982 Molecular Cloning, Cold Spring Harbor Laboratory, N.Y., en las páginas 387 - 389, y especialmente en el párrafo11 que se incorpora aquí como referencia como si estuviera expuesto en su totalidad. Una rigurosidad baja



- se define como si estuviera en 4-6 S X SSC/ 1% (p/v) SDS a 37 - 45 °C durante 2 - 3 horas. Las condiciones de rigurosidad media se considera aquí que son de 1-4X SSC / 0,5% - 1% (p/v) SDS a una temperatura igual o superiores a 45 °C durante 2 - 3 horas. Las condiciones de alta rigurosidad se considera aquí que son de 0,1-1X SSC/0,1% - (% (p/v) SDS a una temperatura igual o superior a 60° C durante 1 - 3 horas. Como se utiliza aquí, el medio para condiciones de alta rigurosidad se refiere a condiciones que son ya sea condiciones de rigurosidad media, condiciones de rigurosidad alta, o condiciones entre rigurosidad media y alta. Preferiblemente, una molécula aislada de ácido nucleico que hibrida con el ORF de los genes *FAO1* o *FAO2* objetivo, hibrida bajo condiciones de alta rigurosidad.
- La presente invención provee además un vector que contiene una secuencia de nucleótidos que codifica a *FAO1* que tiene una secuencia de aminoácidos como la expuesta en la SEQ ID NO: 2. En una modalidad preferida, el vector incluye un gen *FAO1* que tiene una secuencia de nucleótidos como la expuesta en la SEQ ID NO: 1.
- De acuerdo con la presente invención, cualquiera de las secuencias de ácido nucleico descritas aquí anteriormente puede ser incorporada dentro de un vector. Además, debe entenderse que un vector puede incluir también un marco de lectura abierto tanto en el gen *FAO1* como en *FAO2*, así como fragmentos de los mismos.
- Se provee también una célula huésped que es transfectada o transformada con una molécula de ácido nucleico que codifica una *FAO1* y/o *FAO2a* y/o *FAO2b* objetivo. Por ejemplo, una célula huésped puede ser transfectada o transformada con una molécula de ácido nucleico que codifica una secuencia de aminoácidos como la expuesta en cualquiera de las SEQ ID Nos: 2. Preferiblemente, la célula huésped es transformada con una molécula de ácido nucleico que codifica un gen *FAO1* que contiene la secuencia de nucleótidos descrita en la SEQ ID NO: 1. Debe observarse que la presente invención también abarca células huésped transfectadas o transformadas con cualquiera de las secuencias de ácido nucleico descritas aquí anteriormente.
- Dependiendo de la célula huésped transformada con un de *FAO2a* o *FAO2b* objetivo, el codón CUG en el ARNm correspondiente a las posiciones de los nucleótidos 2049 - 2051 de las SEQ ID Nos: 3 y 5 será traducido en forma diferente. Por ejemplo, aunque los codones CUG sean traducidos como serina en *C. tropicalis* otros organismos huésped siguen el código universal y traducen a los codones CUG como leucina. Por lo tanto, con el propósito de garantizar que el codón de ARNm correspondiente a las posiciones de los nucleótidos 2049 - 2051 de la SEQ ID No: 3 y las posiciones de los nucleótidos 1627 - 1629 en la SEQ ID No: 5 en un gen *FAO2* objetivo es traducido en serina en una célula huésped diferente a *C. tropicalis*, las posiciones de los nucleótidos 2049 - 2051 de la SEQ ID No: 3 y las posiciones de los nucleótidos 1627 - 1629 en la SEQ ID No: 5 en un gen *FAO2* objetivo es traducida en serina en una célula huésped diferente a *C. tropicalis*, tal codón debe ser uno que es traducido en serina siguiendo el código universal. Los ejemplos de tales codones incluyen UCU, UCC, UCA, UCG, AGU, y AGC. Los métodos para alterar la secuencia de ADN (y por lo tanto los codones en la correspondiente secuencia transcrita de ARNm) son bien conocidos en el arte e incluyen, por ejemplo, diversas técnicas de mutagénesis *in vitro*. Existen diversos kits comercialmente disponibles particularmente adecuados para esta aplicación tal como el kit de mutagénesis *in vitro* T7-Gen (USB, Cleveland, OH). Alternativamente, se puede emplear tecnología de PCR para alterar la secuencia de nucleótidos. Si un gen *FAO2* es sintetizado químicamente, se puede sintetizar la secuencia de nucleótidos correspondiente a las posiciones 2049 - 2051 de las SEQ ID Nos: 3 y 5 como ADN que transcribe un codón que puede ser traducido en serina de acuerdo con el código universal, por ejemplo, incorporando una secuencia de ADN correspondiente a aquellos codones enlistados anteriormente.
- En una modalidad preferida de la invención, una molécula aislada de ácido nucleico que codifica *FAO1* es manipulada de tal manera que el promotor del gen nativo *FAO1* es removido y se enlaza operativamente un promotor de otro gen con la secuencia que codifica al gen *FAO1*. En forma similar, se manipula preferiblemente una molécula aislada de ácido nucleico que codifica *FAO2* de tal manera que se remueve el promotor del gen nativo *FAO2* y se enlaza operativamente un promotor de otro gen con la secuencia que codifica al gen *FAO2*. El término "operativamente enlazado" se refiere a la asociación de secuencias de ácido nucleico de tal manera que la función de una es afectada por la otra. Un promotor está operativamente enlazado con un marco de lectura abierto de un gen, cuando es capaz de afectar la expresión del marco de lectura abierto (ORF) (es decir, el ORF está bajo el control transcripcional del promotor). A pesar de la presencia de otras secuencias entre el promotor y el ORF, debe entenderse que un promotor puede aún estar operativamente enlazado con el ORF.
- Los promotores deseables para sustitución dentro de un gen *FAO1* o *FAO2* incluyen a los promotores que pueden ser inducidos en diferentes momentos durante la bioconversión en respuesta a ciertos estímulos (por ejemplo, estrés, sustrato, muerte celular) conduciendo por lo tanto a una mayor producción de aldehído y a una mayor producción de ácido dicarboxílico en momentos definidos durante el bioproceso. Los ejemplos de promotores adecuados para enlazar operativamente con los marcos de lectura abiertos de *FAO1* o *FAO2* incluyen por ejemplo, *CYP52A2A*, *CYP52A5A*, y *CYP52A1A* (ver la patente de los Estados Unidos No. 6.331.420, cuya divulgación se incorpora aquí como referencia como si estuviera expuesta completamente). Con respecto al uso del promotor *CYP52A2A*, ver también la solicitud de patente en trámite junto con la presente de Serial No. 09/911.781, cuya divulgación se incorpora aquí como referencia como si estuviera expuesta completamente. El gen *CYP52A2A* de *C.*

*tropicalis* 20336 es un gen de una familia de genes involucrados en el metabolismo del ácido oleico para producir ácido oleico dicarboxílico. El nivel de inducción transcripcional de este gen en una fermentación de ácido oleico está muchas veces (>25) por encima de otros miembros de la misma familia. El ejemplo 6 describe aquí la elaboración de fusiones entre el promotor del gen *CYP52A2A* y el ORF de los genes *FAO1* y *FAO2*. Los promotores de los genes para oxidación en  $\beta$  de *Candida* tales como *POX 4* o *POX 5* pueden estar también operativamente enlazados con un marco ORF de un gen *FAO1* o *FAO2*. Preferiblemente, se utiliza un promotor del gen *CYP52A2A* para dirigir la expresión de un gen *FAO1* o *FAO2*.

Por lo tanto, de acuerdo con la presente invención, se proporciona una secuencia de nucleótidos para un marco de lectura abierto (ORF) de un gen que codifica una FAO que tiene una secuencia de aminoácidos como la expuesta en la SEQ ID NO: 2, en donde el ORF está operativamente enlazado con un. Preferiblemente, una secuencia de nucleótidos para un ORF que codifica una FAO operativamente enlazada con un promotor heterólogo incluye ya sea los nucleótidos 1521 - 3635 de la secuencia de nucleótidos expuesta en la SEQ ID NO: 3 o los nucleótidos 1099 - 3213 de la secuencia de nucleótidos expuesta en la SEQ ID NO: 5. Como se utiliza aquí, el término "promotor heterólogo" significa que incluye a un promotor diferente del promotor nativo asociado con una secuencia que codifica una FAO particular. Estas secuencias de nucleótidos pueden ser descritas por lo tanto también como genes quiméricos.

En otro aspecto de la invención, se provee un vector de expresión que contiene una secuencia de nucleótidos que codifica un gen *FAO1* o *FAO2* operativamente enlazado con una secuencia del promotor heterólogo. El término "vector de expresión" es utilizado aquí ampliamente y se pretende que abarque cualquier medio que incluya una molécula de ácido nucleico y que puede ser utilizada para transformar una célula objetivo. Los vectores de expresión abarcan por lo tanto todos los ejemplos de vectores enlistados aquí incluidos los vectores de integración. Los ejemplos de vectores de expresión incluyen, pero no se limitan a, vectores plásmidos, fagémidos, fago, cósmidos, cromosomas artificiales de levadura o de ADN lineal. Los ejemplos de plásmidos incluyen, pero no se limitan a, por ejemplo, plásmidos episomales de levadura o plásmidos de replicación de levadura.

Además, para amplificar los genes *FAO* para mejorar la productividad de diácido y disminuir la formación de ácido graso T-hidroxi, se pueden utilizar los genes *FAO* para crear construcciones desactivadas para romper los genes *FAO* nativos en *Candida tropicalis*. Los métodos para elaborar construcciones desactivadas y utilizar las mismas para romper los genes son bien conocidos en el arte. La ruptura de genes *FAO* nativos en *C. tropicalis* bloquea la ruta del diácido en esa etapa y permite la construcción de  $\forall$ , compuestos T-dihidroxi cuando se utilizan sustratos de alcano y ácidos grasos T-hidroxi cuando se utilizan sustratos de ácido graso.

La producción de un aldehído a partir de un alcohol durante la segunda etapa de la ruta de oxidación en  $\omega$  de ácidos grasos se puede incrementar por: (a) aislamiento de una molécula de ácido nucleico que contiene una secuencia de codificación para un gen *FAO2*; (b) aislamiento de una secuencia del promotor de un gen que transcribe a una velocidad más alta que el gen *FAO2*, (c) enlazamiento operativo de la secuencia del promotor con el marco de lectura abierto (ORF) del gen *FAO2* para crear un gen de fusión; (d) inserción del gen de fusión dentro de un vector de expresión; (e) transformación de una célula huésped con el vector de expresión y (f) cultivo de la célula huésped transformada en un medio que contiene un sustrato orgánico que puede ser biooxidado a un ácido mono o policarboxílico. Preferiblemente, la molécula de ácido nucleico *FAO2* codifica una proteína *FAO2* que contiene una secuencia de aminoácidos como la expuesta en las SEQ ID Nos: 4 ó 6.

La producción de una cetona a partir de un alcohol durante la segunda etapa de la ruta de oxidación en  $\omega$  de ácidos grasos se puede incrementar por: (a) aislamiento de una molécula de ácido nucleico que contiene una secuencia de codificación para un gen *FAO2*; (b) aislamiento de una secuencia del promotor de un gen que transcribe a una velocidad más alta que el gen *FAO2*, (c) enlazamiento operativo de la secuencia del promotor con el marco de lectura abierto (ORF) del gen *FAO2* para crear un gen de fusión; (d) inserción del gen de fusión dentro de un vector de expresión; (e) transformación de una célula huésped con el vector de expresión y (f) cultivo de la célula huésped transformada en un medio que contiene un sustrato orgánico que puede ser biooxidado a un ácido mono o policarboxílico. Preferiblemente, el sustrato orgánico es un 2-alcohol. Preferiblemente, la molécula de ácido nucleico codifica una proteína *FAO2* que contiene una secuencia de aminoácidos como la expuesta en las SEQ ID Nos: 4 ó 6.

La producción de ácido dicarboxílico se puede incrementar por: (a) aislamiento de una molécula de ácido nucleico que contiene una secuencia de codificación para un gen *FAO2*; (b) aislamiento de una secuencia del promotor de un gen que transcribe a una velocidad más alta que el gen *FAO2*, (c) enlazamiento operativo de la secuencia del promotor con el marco de lectura abierto (ORF) del gen *FAO2* para crear un gen de fusión; (d) inserción del gen de fusión dentro de un vector de expresión; (e) transformación de una célula huésped con el vector de expresión y (f) cultivo de la célula huésped transformada en un medio que contiene un sustrato orgánico que puede ser biooxidado a un ácido mono o policarboxílico. Preferiblemente, el ORF del gen *FAO2* incluye los nucleótidos 1521 - 3635 de una secuencia de nucleótidos como la expuesta en la SEQ ID NO: 3 o los nucleótidos 1099 - 3213 de una secuencia de nucleótidos como la expuesta en la SEQ ID NO: 5.

Se entiende que, en cualquiera de los métodos divulgados aquí, FAO1 y FAO2 pueden ser producidas separadamente en una célula huésped. Alternativamente, tanto FAO1 como FAO2 pueden ser producidas en la misma célula huésped. De este modo, por ejemplo, se puede transformar una célula huésped con: un vector de expresión que contiene un gen *FAO1* o un fragmento del mismo, un vector de expresión que contiene un gen *FAO2a* o un fragmento del mismo, un vector de expresión que contiene un gen *FAO2b* o un fragmento del mismo, un vector de expresión que contiene tanto un *FAO1* y/o un vector de expresión que contiene un gen *FAO1* o un fragmento del mismo, un vector de expresión que contiene un gen *FAO2a* o un fragmento del mismo, un vector de expresión que contiene un gen *FAO2b* o un fragmento del mismo, un vector de expresión que contiene tanto un *FAO1* y/o un gen *FAO2a* o un fragmento del mismo y/o un gen *FAO2b* o un fragmento(s) del mismo, o cualquier combinación de tales vectores.

Debe entenderse que las células huésped dentro de las cuales se han introducido una o más copias de los genes *FAO1* y/o *FAO2* deseados (incluyendo genes quiméricos *FAO* que contienen un gen *FAO1* o *FAO2* operativamente enlazados con un promotor heterólogo) pueden ser elaboradas para incluir tales genes por medio de cualquier técnica conocida por aquellos capacitados en el arte. Por ejemplo, las células huésped adecuadas incluyen procariontes tales como *Bacillus sp.*, *Pseudomonas sp.*, *Actinomycetes sp.*, *Eschericia sp.*, *Mycobacterium sp.*, y eucariotas tales como levaduras, algas, células de insectos, células de plantas y hongos filamentosos. Las células huésped adecuadas son preferiblemente células de levadura tales como *Yarrowia*, *Bebaromyces*, *Saccharomyces*, *Schizosaccharomyces*, y *Pichia* y más preferiblemente aquellas del género *Candida*. Las especies preferidas de *Candida* son *tropicalis*, *maltosa*, *apicola*, *paratropicalis*, *albicans*, *cloacae*, *guillemondii*, *intermedia*, *lipolytica*, *parapsilosis* y *zeylenoides*. Los huéspedes particularmente preferidos incluyen cepas de *C. tropicalis* que han sido genéticamente modificadas de tal manera que uno o más de los genes cromosómicos *POX4A*, *POX4B* y ambos genes *POXS* han sido rotos como se describe, por ejemplo, en las patentes de los Estados Unidos Nos. 5.254.466 y 5.620.878, cada una incorporada aquí como referencia como si estuvieras expuestas en su totalidad. Tal ruptura bloquea la ruta de oxidación en  $\beta$ . Los ejemplos de cepas bloqueadas para oxidación en  $\beta$  de *C. tropicalis* incluyen H41, H41B, His1, H45, H43, H53, H534, H534B, H435 y H5343 (ATCC 20962) como se describe en la patente de los Estados Unidos No. 5.254.466 mencionada anteriormente.

Vectores tales como plásmidos, fagémidos, fagos o cósmidos pueden ser utilizados para transformar o para transfectar células huésped adecuadas. Las células huésped pueden también ser transformadas por medio de la introducción dentro de una célula de un(os) vector(es) lineal(es) de ADN que contienen la secuencia del gen deseado. Tal ADN lineal puede ser conveniente cuando se desea evitar la introducción de un ADN no nativo (foráneo) dentro de la célula. Por ejemplo, el ADN que consiste de un gen(es) objetivo deseado(s) flanqueado(s) por secuencias de ADN que son nativas para la célula puede ser introducido en la célula por medio de electroporación, transformación con acetato de litio, esferoplastia y similares. Las secuencias de flanqueo de ADN pueden incluir marcadores seleccionables y/o otras herramientas para ingeniería genética.

Un sustrato orgánico adecuado para uso en los métodos descritos aquí puede ser cualquier compuesto orgánico que pueda ser biooxidado hasta un ácido mono o policarboxílico (o cualquier compuesto que pueda ser biooxidado hasta un grupo cetona por los métodos descritos aquí dirigidos a la producción de una cetona). Tal compuesto puede ser cualquier compuesto alifático saturado o insaturado o cualquier compuesto aromático carboxcíclico o heterocíclico que tenga al menos un grupo terminal metilo, un grupo terminal carboxilo y/o un grupo terminal funcional que pueda ser oxidado hasta un grupo carboxilo por medio de biooxidación. Un grupo terminal funcional que es un derivado de un grupo carboxilo puede estar presente en la molécula sustrato y puede ser convertido en un grupo carboxilo por medio de una reacción diferente a la biooxidación. Por ejemplo, si el grupo terminal es un éster que no es de *C. tropicalis* de tipo silvestre ni las modificaciones genéticas descritas aquí permitirá la hidrólisis de la función éster hasta un grupo carboxilo, entonces se puede añadir una lipasa durante la etapa de fermentación para liberar ácidos grasos libres. Los sustratos orgánicos adecuados incluyen, pero no se limitan a, ácidos grasos saturados, ácidos grasos insaturados, alcanos, alquenos, alquinos y combinaciones de los mismos.

Los alcanos son un tipo útil aquí de sustrato orgánico saturado. Los alcanos pueden ser de cadena recta o ramificada, lineal o cíclica, sustituidos o no sustituidos. Los alcanos particularmente preferidos son aquellos que tienen aproximadamente desde 4 hasta aproximadamente 25 átomos de carbono, ejemplos de los cuales incluyen, pero no se limitan a, butano, hexano, octano, nonano, dodecano, tridecano, tetradecano, octadecano y similares.

Los ejemplos de sustratos orgánicos no saturados que pueden ser utilizados aquí incluyen, pero no se limitan a, olefinas internas tales como 2-penteno, 2-hexeno, 3-hexeno, 9-octadeceno y similares; olefinas alfa tales como 1-dodeceno, 1-octadeceno, 1-tetradeceno, y similares; ácidos carboxílicos no saturados tales como ácido 2-hexenóico y ésteres del mismo, ácido oleico y ésteres del mismo incluyendo trigliceril ésteres que tienen un contenido relativamente alto de ácido oleico, ácido erúico y ésteres de los mismos incluyendo trigliceril ésteres que tienen un contenido relativamente alto de ácido erúico, ácido ricinoleico y ésteres del mismo incluyendo trigliceril ésteres que tienen un contenido relativamente alto de ácido linoleico, ácido linoleico y ésteres del mismo incluyendo trigliceril ésteres que tienen un contenido relativamente alto de ácido linoleico; alcoholes insaturados tales como 3-hexen-1-ol, 9-octadecen-1-ol, alcoholes saturados tales como 2-decanol, 2-undecanol, 2-dodecanol, 2-hexadecanol, 10-

undecen-1-ol; 1,2-octanodiol; 1,10-decanodiol; 1,2-dodecanodiol; 1,16-hexadecanodiol; ácido 10-hidroxicanoico; ácido 12-hidroxicanoico; ácido 16-hidroxicanoico y similares; aldehídos no saturados tales como 3-hexen-1-al, 9-octadecen-1-al y similares. Además de lo anterior, un sustrato orgánico que puede ser utilizado aquí incluye compuestos alicíclicos que tienen al menos un doble enlace carbono-carbono interno y al menos un grupo terminal etilo, y/o un grupo terminal funcional que puede ser oxidado hasta un grupo carboxilo por medio de biooxidación.

El sustrato orgánico puede contener también otros grupos funcionales que pueden ser biooxidados hasta grupos carboxilo tales como un grupo aldehído o un grupo alcohol. El sustrato orgánico puede contener también otros grupos funcionales que no puedan ser biooxidados hasta grupos carboxilo y no interfieran con la biooxidación tales como halógenos, éteres, y similares.

Los ejemplos de ácidos grasos saturados que pueden ser aplicados a las células que incorporan los presentes genes *FAO1* y *FAO2* incluyen a los ácidos caproico, enántico, caprílico, pelargónico, cáprico, undéclico, láurico, mirístico, pentadecanoico, palmítico, margárico, esteárico, araquídico, behénico y combinaciones de los mismos. Los ejemplos de ácidos grasos insaturados que pueden ser aplicados a células que incorporan los presentes genes *FAO1* y *FAO2* incluyen ácidos palmitoleico, oleico, erúcido, linoleico, linoléico y combinaciones de los mismos. Se pueden aplicar alcanos y fracciones de alcanos que incluyen enlaces de cadena de C12 hasta C24 en cualquier combinación. Un ejemplo de una mezcla preferida de ácidos grasos es el ácido graso de girasol con alto contenido de ácido oleico (HOSFFA por sus siglas en inglés). HOSFFA es una mezcla de ácidos grasos que contiene aproximadamente 80% de ácido oleico y se encuentra comercialmente disponible de Cognis Corporation como Edenor®. Emersol® es otro HOSFFA comercialmente disponible de Cognis Corporation.

La invención es ilustrada adicionalmente por medio de los siguientes ejemplos específicos que no pretenden limitar de ninguna manera el alcance de la invención.

#### EJEMPLO 1

##### Materiales y Métodos

Transformación de *C. tropicalis* utilizando acetato de litio

Se utilizó el siguiente protocolo para transformar *C. tropicalis* de acuerdo con los procedimientos descritos en Current Protocols in Molecular Biology, Supplement 5, 13.7.1 (1989) Frederick M. Ausubel, Roger Brent, Robert E. Kingston, David D. Moore, J. G. Seidman, John A. Smith, and Kevin Struhl, eds., John Wiley and Sons, Hoboken, NJ. Se inocularon 5 ml de YEPD con *C. tropicalis* H5343 ura- y se incubó durante la noche sobre un agitador New Brunswick a 30 °C y 170 rpm. Al día siguiente, se utilizó el cultivo de la noche anterior para inocular 50 ml de YEPD a una OD<sub>600</sub> de 0,2 y se continuó el cultivo a 30 °C, 170 rpm. Se recolectaron las células a una OD<sub>600</sub> de 1,0. Se transfirió el cultivo a un tubo de polipropileno de 50 ml y se centrifugó a 1000 X g durante 10 min. Se resuspendió el precipitado de células en 10 ml de TE estéril (Tris-Cl 10 mM y EDTA 1 mM, pH 8,0). Se centrifugaron nuevamente las células a 1000 X g durante 10 min y se resuspendió el precipitado celular en 10 ml de una solución estéril de acetato de litio (LiAc (acetato de litio 0,1 M, Tris-Cl 10 mM, pH 8,0, EDTA 1 mM). Después de centrifugación a 1000 X g durante 10 min., se resuspendió el precipitado en 0,5 ml de LiAc. Se incubó esta solución durante 1 h a 30°C mientras se agitaba suavemente a 50 rpm. Se incubó una alícuota de 0,1 ml de esta suspensión con 15 - 20 µg de ADN de transformación a 30°C sin agitación durante 30 min. Se añadió e incubó una solución de 0,7 ml de PEG (40 % en peso/vol. de polietilén glicol 3340, acetato de litio 0,1 M, Tris-Cl 10 mM, pH 8,0, EDTA 1 mM) a 30°C durante 45 min. Se colocaron luego los tubos a 42°C durante 10 min. Se sembró en placa una alícuota de 0,2 ml sobre medio completo sintético menos uracilo (SC - uracilo) (Kaiser et al. Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press, EUA, 1994). Se monitoreó el crecimiento de los transformantes durante 5 días. Después de tres días, se recolectaron varios transformantes y se los transfirió a placas de SC-uracilo para la preparación y selección del ADN genómico.

Fermentaciones

Las fermentaciones fueron realizadas utilizando ácido graso de girasol con alto contenido de ácido oleico (HOSFFA como el sustrato, utilizando glucosa como el co-sustrato). El HOSFFA utilizado era aproximadamente 85% de ácido oleico (84,45% de ácido oleico, 5,24% de ácido linoleico, 4,73% de ácido esteárico y 3,87% de ácido palmítico, 1,71% de otros ácidos grasos). El procedimiento de fermentación tenía tanto una fase de crecimiento como una de bioconversión. Se cultivaron las células sobre glucosa a 35 °C hasta una absorbancia a 600 nm de 50 a 60. Cuando se completó la fase de crecimiento de acuerdo a lo determinado por medio de un fuerte aumento en oxígeno disuelto, las células fueron inmediatamente pasadas a la fase de bioconversión, la cual fue realizada a una temperatura de 30 °C. Se utilizó una alimentación lenta de glucosa (0,87 g/L/h desde 0 hasta 42 h y 0,7 g/L/h después de eso) para suministrar energía. Al mismo tiempo se añadió una pequeña carga de HOSFFA (0,55% v/v) a la fermentación. Se continuó la alimentación de sustrato (velocidad promedio de alimentación de 1,6 g/L/h) durante el resto de la fermentación. H5343 es la cepa base derivada de *Candida tropicalis* (ATCC 20336) bloqueada para

oxidación en  $\beta$ , que ha tenido los genes *POX4* y *POX5* rotos por medio de mutagénesis de inserción. La cepa HDC23-3 es una cepa base de H5343 que es amplificada por un gen para la citocromo P450 monooxigenasa, *CYP52A2*, y el gen para la citocromo P450 reductasa (NCP), habiendo sido clonados ambos genes a partir de *C. tropicalis* ATCC 20336.

## 5 Perfil de actividad de FAO durante fermentaciones con HOSFFA

### Lavado de células

Se prepararon y ensayaron células HDC23-3 a partir de diferentes procedimientos de fermentación, y tomadas en diferentes momentos durante los procesos de fermentación, por la actividad de FAO. Debido a los altos niveles de diácido sólido en el caldo, particularmente en muestras tomadas después en la fermentación, las muestras tuvieron que ser lavadas extensivamente para remover el diácido antes de elaborar los extractos. Ya que se encontró que las muestras congeladas pierden una cantidad de actividad enzimática (después del lavado y preparación de los extractos), se colocó caldo fresco de fermentación sobre hielo hasta que estuviera listo para lavado. Se mezcló perfectamente esta muestra y se removió una muestra de 20 ml. Se centrifugó esta muestra (centrífuga refrigerada Sorvall RT6000B) aproximadamente a 1.500 x g (rotor H1000B a 2500 rpm) durante 5 minutos para precipitar las células. Se decantó el sobrenadante, y se resuspendieron las células en 40 ml de amortiguador HEPES 50 mM, pH 7,6. Se centrifugó la muestra a 1.500 x g únicamente durante un minuto, que precipitó las células de levadura, pero permitió que el producto diácido menos denso "flotara" en el sobrenadante. Se decantó el sobrenadante y se lavaron las células nuevamente en 40 ml de amortiguador. Después de la segunda etapa de lavado, se examinó el precipitado celular por cualquier signo de diácido residual, evidenciado por un precipitado de color blanco en la parte superior del precipitado celular, o parches de color blanco en el mismo precipitado celular. Se repitió el lavado hasta que el precipitado estaba completamente libre de todo el diácido y el precipitado celular era completamente de un color café claro. Durante el transcurso de la fermentación, a medida que las células producían más diácido, fueron necesarias más etapas de lavado para liberar al precipitado celular del diácido.

### Preparación de extractos libres de células

Después de retirar por medio de lavado el diácido, se resuspendió el precipitado celular en 10 ml de amortiguador de fosfato de potasio 50 mM con 20% de glicerol, pH 7,6 (amortiguador de fosfato-glicerol) para una concentración dos veces mayor de enzimas celulares. Se añadieron 100  $\mu$ l de una solución 100 mM del inhibidor de serina-proteasa, fluoruro de fenilmetilsulfonilo (PMSF por sus siglas en inglés) en isopropanol, para una concentración final de 1 mM. Se rompieron luego las células pasando la muestra tres veces a través de una celda a presión francesa refrigerada (SLM Instruments, Inc. French Pressure Cell Press) aproximadamente a 20.000 psig. Se almacenó la muestra sobre hielo antes, durante y después del rompimiento de las células para evitar la pérdida de actividad de la enzima. Esta suspensión de células rotas fue luego centrifugada aproximadamente a 37,000 x g durante 30 minutos para precipitar los residuos celulares, dejando las enzimas celulares en el sobrenadante. Se removió el sobrenadante (extracto libre de células) y guardó para llevar a cabo los ensayos de la enzima. Durante los ensayos, se mantuvo siempre sobre hielo este extracto celular, y fue almacenado a -20°C entre los grupos de ensayos para preservar la actividad de la enzima.

### Ensayo de enzimas FAO

El procedimiento del ensayo utilizado es una modificación de un ensayo utilizado por Kemp et al (1). El ensayo es una reacción acoplada de doble enzima. Se utilizó dodecanol como el sustrato para la FAO, que oxida el dodecanol hasta dodecanal. Al mismo tiempo este reduce el oxígeno molecular hasta peróxido de hidrógeno. Una segunda enzima en la mezcla de reacción, peroxidasa derivada de rábano picante (HRP), utiliza electrones obtenidos por oxidación con peróxido de hidrógeno para reducir al ácido 2,2'-azino-bis(3-etilbenzotiazolina-6-sulfónico) (ABTS) en la mezcla de reacción. El ABTS reducido absorbe fuertemente a 405 nm. La calidad de la HRP es importante, ya que se ha observado que las preparaciones obtenidas comercialmente varían en calidad. Es necesario, por lo tanto, obtener preparaciones de alta pureza. La HRP (Sigma #P8415) fue elaborada en una concentración de 2 mg/ml ( $\approx$ 250 unidades/mg) en amortiguador de fosfato de potasio 50 mM, pH 7.6. En experimentos en los cuales se varió la cantidad de peroxidasa de rábano en la mezcla de reacción, se encontró que 5  $\mu$ l de esta solución en una mezcla de reacción de un ml era suficiente para obtener velocidad máxima.

La mezcla final de reacción (1,0 ml) para el ensayo general de alcohol oxidasa de: 500  $\mu$ l de amortiguador HEPES 200 mM, pH 7,6; 50  $\mu$ l de una solución de ABTS de 10 mg/ml en agua desionizada; 10  $\mu$ l de una solución 5 mM de dodecanol in acetona; y 5  $\mu$ l de una solución de peroxidasa de rábano de 2 mg/ml en amortiguador de fosfato de potasio 50 mM, pH 7,6. Después de añadir el extracto, se leyó generalmente la actividad a 405 nm durante un minuto a temperatura ambiente. La cantidad de extracto que se añadió a la mezcla de reacción fue variada de tal manera que la actividad cayó dentro del rango de 0,2 a 1,0  $\Delta A_{405 \text{ nm}}/\text{min}$ . Se reportó la actividad de la alcohol oxidasa como unidades de actividad específica/mg de proteína (1 unidad =  $\mu$ moles de sustrato oxidado/min). Se utilizó un coeficiente de extinción a 405 nm de 18,4 para ABTS y era equivalente a sustrato oxidado 0,5 mM. En

ciertos experimentos de especificidad de sustrato, se utilizó amortiguador HEPES 200 mM, pH 7,6, que contenía 0,5% de Triton X100 en lugar del amortiguador HEPES 200 mM en la mezcla de reacción descrita anteriormente. El detergente ayudó en la solubilización de algunos de los sustratos más insolubles en agua analizados.

#### Preparación de microsomas

- 5 Los microsomas fueron preparados a partir de extractos libres de células de *Candida tropicalis* por medio de centrifugación a 100.000 x g durante una hora a 4°C. Se encontró la FAO dentro de los microsomas precipitados. La catalasa, una enzima soluble, permaneció en el sobrenadante, lo que permitió la separación de estas dos enzimas. El sobrenadante fue removido y analizado por la actividad de FAO y de la catalasa. Se resuspendió el precipitado microsomal en una cantidad equivalente de amortiguador de fosfato-glicerol, después de lo cual también fue  
10 analizado por la actividad de la catalasa y de FAO. Se prepararon microsomas de *Escherichia coli* en la misma forma.

#### Determinación de proteína

- 15 Se determinó la concentración de proteína en los extractos. El método de Lowry para determinación de concentración de proteína produjo resultados mucho más altos de lo esperado, y fueron inconsistentes entre las determinaciones de proteína de las mismas muestras. Por esta razón, se llevó a cabo un nuevo procedimiento utilizando el Reactivo de Bradford (Sigma, #B6916) y siguiendo el protocolo suministrado por el proveedor.

## EJEMPLO 2

### Clonación de genes para la oxidasa de alcohol graso

#### Preparación de ADN genómico

- 20 Se inocularon 50 ml de caldo YPD (Difco) con una sola colonia de *C. tropicalis* 20336 a partir de una placa de agar YPD (Difco) y se desarrolló durante la noche a 30°C. Se inocularon 5 ml del cultivo durante la noche en 100 ml de caldo YPD fresco y se incubó a 30°C durante 4 a 5 h con agitación. Se recolectaron las células por medio de centrifugación, se lavó dos veces con agua destilada estéril y se resuspendió en 4 ml d amortiguador de esferoplastia (Sorbitol 1 M, EDTA 50 mM, mercaptoetanol 14 mM) y se incubó durante 30 min a 37°C con agitación  
25 suave. Se añadieron 0,5 ml de zimoliasa de 2 mg/ml (ICN Pharmaceuticals, Inc., Irvine, CA) y se incubó a 37°C con agitación suave durante 30 a 60 min. Se monitoreó la formación de esferoplastos por medio de lisis con SDS. Se recolectaron los esferoplastos por medio de una breve centrifugación (4.000 rpm, 3 min) y fueron lavados una vez con el amortiguador de esferoplastos sin mercaptoetanol. Los esferoplastos recolectados fueron luego suspendidos en 4 ml de amortiguador de lisis (Tris 0,2 M /pH 8.0, EDTA 50 mM, 1% de SDS) que contenía 100 mg/ml de ARNasa (Qiagen Inc., Chatsworth, CA) y se incubó a 37°C durante 30 a 60 min.

- 35 Se desnaturalizaron y extrajeron las proteínas dos veces con un volumen igual de cloroformo/alcohol isoamílico (24:1) mezclando suavemente las dos fases por medio de inversión manual. Se separaron las dos fases por centrifugación a 10.000 rpm durante 10 min y se recuperó la fase acuosa que contenía el ADN de alto peso molecular. Se añadió NaCl a la capa acuosa hasta una concentración final de 0,2 M y se precipitó el ADN por medio de la adición de 2 volúmenes de etanol. Se enrolló el ADN precipitado con una varilla de vidrio limpia y se resuspendió en amortiguador TE (Tris 10 mM /pH 8.0, EDTA 1 mM) y se permitió que se disolviera durante la noche a 4°C. Al ADN disuelto, se le añadió ARNasa libre de cualquier actividad de ADNasa (Qiagen Inc., Chatsworth, CA) hasta una concentración final de 50 mg/ml y se incubó a 37°C durante 30 min. Se añadió luego la proteasa (Qiagen Inc., Chatsworth, CA) hasta una concentración final de 100 mg/ml y se incubó a 55 - 60°C durante 30 min. Se extrajo  
40 la solución una vez con un volumen igual de fenol/cloroformo/alcohol isoamílico (25:24:1) y una vez con un volumen igual de cloroformo/alcohol isoamílico (24:1). A la fase acuosa se le añadieron 0,1 volúmenes de acetato de sodio 3 M y 2 volúmenes de etanol enfriado con hielo (absoluto) y se enrolló el ADN de alto peso molecular con una varilla de vidrio y se disolvió en 1 a 2 ml de amortiguador TE.

#### Preparación de la biblioteca

- 45 Se construyó una biblioteca genómica utilizando el vector  $\lambda$  ZAP Express™ (Stratagene, La Jolla, CA). Se digirió parcialmente ADN genómico con Sau3A1 y se purificaron fragmentos en el rango de 6 a 12 kb a partir de un gel de agarosa después de electroforesis del ADN digerido. Estos fragmentos de ADN fueron luego ligados a los brazos del vector  $\lambda$  ZAP Express™ digerido con *Bam*HI de acuerdo con los protocolos del fabricante. Se establecieron tres ligaciones para obtener aproximadamente  $9,8 \times 10^5$  clones independientes. Todas las tres bibliotecas fueron  
50 reunidas y amplificadas de acuerdo con las instrucciones del fabricante para obtener un patrón de título alto ( $>10^9$  unidades formadoras de placa/ml) para almacenamiento durante un largo tiempo. El título de la biblioteca empacada de fagos fue comprobada después de la infección de células XL1Blue-MRF<sup>1</sup> de *E. coli*. Se cultivaron las células

5 XL1Blue-MRF' de *E. coli* durante la noche ya sea en medio LB (Difco) o NZCYM (Difco) que contenía MgSO<sub>4</sub> 10 mM y 0,2% de maltosa a 37°C ó 30°C, respectivamente con agitación. Las células fueron luego centrifugadas y resuspendidas en 0,5 a 1 volúmenes de MgSO<sub>4</sub> 10 mM. 200 µl de este cultivo de *E. coli* fueron mezclados con diferentes diluciones de la biblioteca empacada de fagos e incubados a 37°C durante 15 min. A esta mezcla se le añaden 2,5 ml de agarosa de superficie LB o agarosa de superficie NZCYM (mantenida a 60°C) y se siembran en placa sobre agar LB (Difco) o agar NCZYM (Difco) presente en cajas Petri de 82 mm. Se permitió que se propagaran los fagos durante la noche a 37 °C para obtener placas discretas y se determinó el título del fago.

Selección de bibliotecas genómicas (Formadoras de placa)

Siembra en placas de la biblioteca λ

10 Se cultivaron células XL1Blue-MRF' de *E. coli* durante la noche en medio LB (25 ml) que contenía MgSO<sub>4</sub> 10 mM y 0,2% de maltosa a 37 °C, 250 rpm. Se centrifugaron luego las células (2200 x g durante 10 min) y se las resuspendió en 0,5 volúmenes de MgSO<sub>4</sub> 10 mM. Se mezclaron 500 µl de este cultivo de *E. coli* con una suspensión de fagos que contenía 25.000 partículas de fagos lambda amplificados y se incubó a 37°C durante 15 min. A esta mezcla se le añadieron 6,5 ml de agarosa de superficie NZCYM (mantenida a 60°C) y se sembraron en placa sobre  
15 80 - 100 ml de agar NCZYM presente en una caja de Petri de 150 mm. Se permitió que se propagaran los fagos durante la noche a 37 °C para obtener placas discretas. Después de crecimiento durante la noche, se almacenaron las placas en un refrigerador durante 1 - 2 h antes de recoger las placas.

Recolección de las placas

20 Se colocaron membranas de nylon Magna Lift™ (Micron Separations, Inc., Westbrough, MA) sobre la superficie de agar en contacto completo con las placas y se permitió que se llevara a cabo la transferencia de las placas a membranas de nylon durante 5 min a temperatura ambiente (RT). Después de la transferencia de las placas se colocó la membrana sobre 2 hojas de papel filtro Whatman 3M™ (Whatman, Hillsboro, OR) saturado con una solución de NaOH 0,5 N, NaCl 1,0 M y se dejó durante 10 min a temperatura ambiente (RT) para desnaturalizar el ADN. Se removió el exceso de solución de desnaturalización por medio de transferencia brevemente sobre papel  
25 Whatman 3M™ seco. Se transfirieron luego las membranas a 2 hojas de papel Whatman 3M™ saturado con Tris-HCl 0,5 M (pH 8,0), NaCl 1,5 M y se dejó durante 5 min hasta neutralización. Se lavaron luego las membranas durante 5 min en 200 - 500 ml de 2 X SSC, se secó al aire y luego en horno durante 30 - 40 min a 80°C. Las membranas fueron luego sondadas con ADN marcado.

Hibridaciones de ADN

30 Se lavaron previamente las membranas con una 200 - 500 ml de una solución de 5 X SSC, 0,5% de SDS, EDTA 1 mM (pH 8,0) durante 1 - 2 h a 42°C con agitación (60 rpm) para remover los restos de las bacterias de las membranas. Se hibridaron previamente las membranas durante 1 - 2 h a 42°C (en un volumen equivalente a 0,125 - 0,25 ml/cm<sup>2</sup> de membrana) con amortiguador ECL Gold™ (Amersham) que contiene NaCl 0,5 M y 5% de reactivo de bloqueo. Lo fragmentos de ADN utilizados como sondas fueron purificados a partir del gel de agarosa utilizando un  
35 kit de extracción del gel QIAEX II™ (Qiagen Inc., Chatsworth, CA) de acuerdo con el protocolo del fabricante y marcados utilizando un kit de marcación directa de ácido nucleico ECL™ de Amersham (Amersham). Se añadió ADN marcado (5 - 10 ng/ml de solución de hibridación) a las membranas previamente hibridadas y se permitió que procediera la hibridación durante la noche. Al día siguiente se lavaron las membranas con agitación (60 rpm) dos veces a 42°C durante 20 min cada vez en (en un volumen equivalente a 2 ml/cm<sup>2</sup> de membrana) un amortiguador  
40 que contenía ya sea 0,1 (alta rigurosidad) o 0,5 (baja rigurosidad) X SSC, 0,4% de SDS y 360 g/l de urea. Esto fue seguido por dos lavados de 5 min a temperatura ambiente en (en un volumen equivalente a 2 ml/cm<sup>2</sup> de membrana) 2 X SSC. Se generaron las señales utilizando el reactivo de detección de ácido nucleico ECL™ y se hizo la detección utilizando Hyperfilm ECL™ (Amersham).

45 Se tomaron tapones de agar que contenían placas correspondientes a señales positivas sobre la película de rayos X de las placas patrón utilizando el extremo abierto de una pipeta estéril Pasteur. Se seleccionaron las plaquetas alineando las placas con la película de rayos X. En esta etapa, se tomaron generalmente múltiples plaquetas. Se eluyeron las partículas de fago de los tapones de agar remojando en 1 ml de amortiguador SM (Sambrook et al., (1989) (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.) durante la noche. El fago eluido fue luego diluido y sembrado en placa con células cultivadas en forma fresca  
50 XL1Blue-MRF' de *E. coli* para obtener 100 - 500 plaquetas por placa de agar NCZYM de 85 mm. Se transfirieron las plaquetas a membranas de nylon Magna Lift como anteriormente y sondadas nuevamente utilizando la misma sonda. Se recolectaron las plaquetas aisladas de un solo pozo correspondientes a señales sobre la película de rayos X por remoción de los tapones de agar y eluyendo el fago por remojo durante la noche en 0,5 ml de amortiguador SM.

55 Conversión de clones λ en un plásmido

Para convertir el vector  $\lambda$  ZAP Express™ en un plásmido, se utilizaron cepas de *E. coli* XLBlue-MRF' y XLOR. La conversión se realizó de acuerdo con los protocolos del fabricante (Stratagene) por escisión de una sola plaqueta.

Generación de una sonda para biblioteca

Selección del iniciador

- 5 La sonda para la biblioteca fue preparada a partir de un fragmento de reacción en cadena de polimerasa (PCR) del genoma ATCC 20336 de *C. tropicalis* que se sabe que representa al gen o genes *FAO*. Por medio de la comparación de regiones de homología entre dos genes *FAO* publicados de *C. cloacae* y un gen *FAOT* publicado de *C. tropicalis* (NCYC 470) (4), se desarrollaron dos iniciadores no degenerativos por la PCR: dos iniciadores hacia adelante denominados como (*FAO-F1*) y (*FAO-F2*) y un iniciador inverso denominado como (*FAO-R1*) (ver la Tabla I). Estos iniciadores fueron anticipados para amplificar una región en el gen en el marco de lectura abierto (ORF) del gen *FAO*.

Tabla 1. Comparación de la secuencia del iniciador con la secuencia correspondiente en *Candida cloacae*.

Gen	Comparación de la secuencia	Iniciador
C. t. FAOT	352 5' TCG TGG CGT GAC TCT CCT 3' 369 (SEQ ID NO: 35)	FAO-F1
C. c. FAO1	340 5' TCA TGG AGA GAC TCT CCT 3' 357 (SEQ ID NO: 36)	
C. c. FAO2	340 5' TCA TGG AGA GAC TCT CCA 3' 357 (SEQ ID NO: 37)	
C. t. FAOT	608 5' CTG GTG CTG GTG TAG T 3' 623 (SEQ ID NO:38)	FAO-F2
C. c. FAO1	593 5' CAG GAG CAG GTG TGG T 3' 608 (SEQ ID NO:39)	
C. c. FAO2	593 5' CGG GAG CAG GAG TGG T 3' 608 (SEQ ID NO:40)	
C. t. FAOT	1780 5' TTG GTA CCC ATG CTT GTG G 3' 1762 (SEQ ID NO:41)	FAO-R1
C. c. FAO1	1762 5' TTG GTA CCC AAG CTT GTG G 3' 1744 (SEQ ID NO:42)	
C. c. FAO2	1762 5' TTG GTA CCC AAG CTT GTA G 3' 1744 (SEQ ID NO:43)	

Condiciones de la PCR

- 15 Los iniciadores seleccionados fueron utilizados en una reacción PCR con las siguientes condiciones: 5  $\mu$ l de amortiguador 10X PCR, 5  $\mu$ l ya sea del patrón del iniciador *FAO-F1* o de *FAO-F2* (20  $\mu$ M), 5  $\mu$ l del patrón del iniciador 30R1 (20  $\mu$ M), 1  $\mu$ l de la mezcla de nucleótidos, 0,5  $\mu$ l de Taq polimerasa, y 1  $\mu$ l de ADN genómico de la cepa ATCC 20336 de *Candida tropicalis*, todo en un volumen de reacción de 50  $\mu$ l. Las condiciones de reacción fueron: 95 °C durante 2 min, seguido por 35 ciclos de 95 °C durante 30 s, 53 °C durante 30 s y 72 °C durante 1 min.
- 20 El par de iniciadores, *FAO-F1* y *FAO-R1*, que deben haber producido un fragmento esperado de 1429 pb, nunca produjeron un producto PCR. El par de iniciadores, *FAO-F2* y *FAO-R1*, que deben haber producido un fragmento esperado de 1173 pb, generaron un fragmento de aproximadamente 1200 pb de longitud.

Clonación del producto PCR con TOPO TA

- 25 El producto PCR fue clonado utilizando un kit de clonación TOPO TA (Invitrogen, Carlsbad, Ca., TOPO TA Cloning Kit, Versión K2, #25-0184) en células de la cepa Top10F'. El producto PCR fue primero purificado por electroforesis en agarosa de bajo punto de fusión (1%) seguido por escisión de la banda apropiada. La rebanada de gel fue colocada en una tubo de una micrófuga e incubada a 65 °C hasta que fundió el gel, después de lo cual fue colocada a 37 °C. En un tubo nuevo se combinaron 4  $\mu$ l de la agarosa fundida que contenía al producto PCR con 1  $\mu$ l del vector de clonación TOPO TA y se incubó a 37 °C durante 10 minutos. Se logró la transformación mezclando 2 - 4  $\mu$ l



de esta mezcla de reacción con un vial (50 µl) de células competentes de *E. coli* suministradas con el kit. Estas fueron incubadas sobre hielo durante 30 min. Las células fueron sometidas a choque térmico a 42 °C durante 30 segundos sin agitación. Se transfirió inmediatamente el vial a hielo y se añadieron 250 µl de medio SOC a temperatura ambiente. Se agitó luego este tubo en forma horizontal a 37 °C durante 30 min. Se esparcieron 5 alícuotas sobre placas de LB + 100 µg/ml de ampicilina después de lo cual se habían esparcido 25 µl de isopropiltiogalactósido (IPTG, 100 mM) + 40 µl de Xgal (5-bromo-4-cloro-3-indolil β-D-galactopiranosido, 40 mg/ml) al menos 30 minutos antes. Se determinaron los transformantes que contenían insertos por selección azul/blanco, indicando el azul una transformación exitosa que contenía un inserto.

Las colonias blancas que eran positivas para la presencia de un inserto fueron inoculadas en medio LB que contenía 100 µg/ml de ampicilina y se desarrollaron durante la noche. Se obtuvo el ADN del plásmido a partir de estos cultivos utilizando el método del kit Qiaprep (Qiagen, Qiaprep Spin Miniprep Kit #27106), y se analizó la presencia del inserto por restricción con *EcoRI*. En el vector PCR 2.1, el inserto está flanqueado por cualquier lado por sitios *EcoRI*, de tal manera que cortando con *EcoRI* se liberará lo que se ha insertado dentro del plásmido. Diferentes clones mostraron 10 el tamaño correcto del inserto de aproximadamente 1200 pb. Dos de los clones fueron secuenciados y la secuencia de ADN fue comparada con la secuencia publicada para *FAOT* (12). Ya que el grado de homología para ambos clones fue muy alto (79% de coincidencia de las bases 608-1199 del ORF de *FAOT*), uno de ellos fue seleccionado para preparar la sonda de ADN.

#### Preparación de la sonda de ADN

Se obtuvieron varios microgramos de ADN del plásmido de un clon positivo y se digirió el ADN con *EcoRI* para liberar al inserto. La electroforesis del digerido sobre 1,2% de un gel de agarosa de bajo punto de fusión permitió la separación del fragmento de *FAO* del vector PCR 2.1. Se extrajo el ADN apropiadamente dimensionado del gel (Qiaprep, QIAquick Gel Extraction Kit, #28706) y fue cuantificado con un fluorómetro. El fragmento de ADN de *FAO* fue luego marcado utilizando el método ECL (Amersham, ECL kit, #RPN 3001).

### EJEMPLO 3

25 Clonación de genes *FAO* de ATCC 20336 de *Candida tropicalis*

Se elaboraron placas de la biblioteca del fago λ de *C. tropicalis* y el levantamiento de estas placas sobre filtros de membrana de nitrocelulosa se llevó a cabo, de acuerdo al procedimiento descrito en Materiales y métodos (Ejemplo 1). Se identificaron clones positivos putativos como se indica en Materiales y Métodos (Ejemplo 1). Se cultivaron las células XL0LR que contienen estos fragmentos de biblioteca y se obtuvo el ADN del plásmido utilizando el kit Qiaprep. La digestión de restricción y los análisis por PCR confirmaron la presencia de un gen *FAO* en estos clones de biblioteca de *C. tropicalis*. Era conocido a partir de la información de la secuencia del ADN de la sonda que al menos algunos de los clones deben cortarse con *PvuII* y *KpnI*. Por lo tanto, los clones de la biblioteca fueron digeridos con *EcoRI*, *PvuII*, y *KpnI* en digestiones simple, y con *PvuII* y *KpnI* en una digestión doble. Esto permitió 30 determinar el direccionamiento del gen *FAO* y estimar su posicionamiento dentro del inserto del vector PBK-CMV. Los iniciadores iniciales utilizados en la preparación de la sonda de ADN, *FAO-F2* y *FAO-R1*, fueron utilizados para selección por PCR los clones de la biblioteca utilizando ADN de plásmido purificado de estos clones como el molde. Se utilizó ADN genómico de ATCC 20336 de *C. tropicalis* como molde para la reacción PCR en el control. Ocho clones de la biblioteca de *FAO*, denominadas A1, A4, A5, A6, A8, A9, B5, y B6, fueron identificados como clones positivos putativos y fueron enviados a la Sequetech Corporation para secuenciación.

40 Cuando las secuencias de ADN de los clones fueron comparadas con la secuencia publicada de *FAOT*, los clones cayeron en dos grupos. El grupo 1 estaba compuesto de los clones A4, A8, B5, y B6. El grupo 2 estaba compuesto de A1, A5, A6, y A9.

4297 pb del gen del clon AS, que fue denominado *FAO1*, fue secuenciado bicatenario (SEQ ID No: 1). Además del marco de lectura abierto (ORF), que tenía 2112 pb de longitud, había un ADN secuenciado de 1940 pb secuencia 45 arriba y 242 pb secuencia abajo. 4158 pb del gene del clon A9, que fue denominado *FAO2a*, fue secuenciado bicatenario (SEQ ID No: 3). Además del marco de lectura abierto (ORF), que tenía 2112 pb de longitud, había un ADN secuenciado de 1520 pb secuencia arriba y 523 pb secuencia abajo. Existe un codón CTG en los pb 2049 - 2051 en la secuencia de *FAO2* presentada en la SEQ ID NO: 3 (pb 529 - 531 del ORF). Este CTG es denominado como leucina en el código universal, pero el análisis de la secuencia (14) ha confirmado que CTG realmente codifica 50 para serina en ATCC 20336 de *Candida tropicalis*.

La secuenciación adicional del clon A6 demostró una cercana homología con el clon A9. Hubo unas pocas diferencias en los pares de bases, sin embargo, se llevó a cabo tal secuenciación bicatenaria del gen. Los resultados (SEQ ID No: 5, Tabla 2 y Tabla 3) mostraron que el clon A6 era más probablemente un alelo del clon A9. Fue denominado *FAO2b*.

Las regiones del ORF de *FAO1*, *FAO2a* y *FAO2b* fueron comparadas con regiones análogas del gen *FAOT* publicado (12) de *Candida tropicalis* (NCYC 470), los genes *FAO1* y *FAO2* de *Candida cloacae* y el gen *FAO* de *Candida albicans* (Tabla 2).

5 Tabla 2. Comparación de la secuencia de ADN entre los genes objetivo *FAO1* o *FAO2* de *C. tropicalis* y los datos de secuencia publicados para los genes NCYC 470 de *C. tropicalis*, *FAO* de *C. cloacae*, y de *C. albicans*

Secuencia #1 de ADN	Secuencia #2 de ADN	% de identidad
<i>FAO1</i> de <i>C. tropicalis</i>	<i>FAO2a</i> de <i>C. tropicalis</i>	82
<i>FAO1</i> de <i>C. tropicalis</i>	<i>FAOT</i> de <i>C. tropicalis</i>	77
<i>FAO1</i> de <i>C. tropicalis</i>	<i>FAO</i> de <i>C. albicans</i>	71
<i>FAO1</i> de <i>C. tropicalis</i>	<i>FAO1</i> de <i>C. cloacae</i>	62
<i>FAO2a</i> de <i>C. tropicalis</i>	<i>FAOT</i> de <i>C. tropicalis</i>	78
<i>FAO2a</i> de <i>C. tropicalis</i>	<i>FAO</i> de <i>C. albicans</i>	73
<i>FAO2a</i> de <i>C. tropicalis</i>	<i>FAO1</i> de <i>C. cloacae</i>	63
<i>FAO2b</i> de <i>C. tropicalis</i>	<i>FAO1</i> de <i>C. tropicalis</i>	81
<i>FAO2b</i> de <i>C. tropicalis</i>	<i>FAO2a</i> de <i>C. tropicalis</i>	95
<i>FAO2b</i> de <i>C. tropicalis</i>	<i>FAOT</i> de <i>C. tropicalis</i>	77
<i>FAO2b</i> de <i>C. tropicalis</i>	<i>FAO</i> de <i>C. albicans</i>	73
<i>FAO2b</i> de <i>C. tropicalis</i>	<i>FAO1</i> de <i>C. cloacae</i>	62
<i>FAOT</i> de <i>C. tropicalis</i>	<i>FAO1</i> de <i>C. cloacae</i>	62
<i>FAOT</i> de <i>C. tropicalis</i>	<i>FAO</i> de <i>C. albicans</i>	73
<i>FAO1</i> de <i>C. cloacae</i>	<i>FAO</i> de <i>C. albicans</i>	59
<i>FAO1</i> de <i>C. cloacae</i>	<i>FAO2</i> de <i>C. cloacae</i>	79

Se compararon también las secuencias de aminoácidos, que se derivaron utilizando el código universal (Tabla 3).

10 Tabla 3. Comparación de la secuencia derivada de aminoácidos entre los genes objetivo *FAO1* o *FAO2* de *C. tropicalis* y los datos de secuencia publicados para los genes NCYC 470 de *C. tropicalis*, *FAO* de *C. cloacae*, y de *C. albicans*

Proteína #1	Proteína #2	% de Identidad	% de Similitud
<i>FAO1</i> de Cognis	<i>FAO2a</i> de Cognis	81	92
<i>FAO1</i> de Cognis	<i>FAO</i> de <i>C. tropicalis</i>	82	90
<i>FAO1</i> de Cognis	<i>FAO</i> de <i>C. albicans</i>	74	88

(continuación)

Proteína #1	Proteína #2	% de Identidad	% de Similitud
FAO1 de Cognis	FAO1 de <i>C. cloacae</i>	60	76
FAO2a de Cognis	FAO de <i>C. tropicalis</i>	85	93
FAO2a de Cognis	FAO de <i>C. albicans</i>	78	88
FAO2a de Cognis	FAO1 de <i>C. cloacae</i>	59	76
FAO2b de Cognis	FAO1 de Cognis	80	91
FAO2b de Cognis	FAO2a de Cognis	97	98
FAO2b de Cognis	FAOT de <i>C. tropicalis</i>	85	93
FAO2b de Cognis	FAO de <i>C. albicans</i>	76	88
FAO2b de Cognis	FAO1 de <i>C. cloacae</i>	59	75
FAO de <i>C. tropicalis</i>	FAO1 de <i>C. cloacae</i>	60	77
FAO de <i>C. tropicalis</i>	FAO de <i>C. albicans</i>	76	87
FAO1 de <i>C. cloacae</i>	FAO de <i>C. albicans</i>	55	72
FAO1 de <i>C. cloacae</i>	FAO2 de <i>C. cloacae</i>	76	88

5 Los genes *FAO2a* y *FAO2b* son 95% idénticos por la secuencia de ADN y tienen 97% de identidad y 98% de similitud por secuencia de aminoácidos.

10 Los genes *FAO1* y *FAO2a* de la presente invención son 82 % idénticos por la secuencia de ADN (81% para *FAO2b*) y tienen 81 % de identidad y 92 % de similitud por secuencia de aminoácidos (80% de identidad y 91 % de similitud para *FAO2b*). En comparación, los genes *FAO1* y *FAO2* de *C. cloacae* son 79 % idénticos por secuencia de ADN y tienen 76 % de identidad y 88% de similitud por secuencia de aminoácidos. Estos resultados indican que la igual que *C. cloacae*, la cepa 20336 de *C. tropicalis* tiene dos diferentes genes para la oxidasa de alcohol graso.

15 En forma muy interesante, Vanhanen et al. (4) identificaron únicamente un gen *FAO* en su cepa NCYC 470 de *C. tropicalis*. La secuencia de ADN de *FAOT* era 77 % idéntica a los genes *FAO1* y *FAO2b* de la presente invención y 78% idéntica al gen *FAO2a* de la presente invención. La comparación de la secuencia de aminoácidos mostró que *FAOT* tenía 82 % de identidad y 92 % de similitud con el gen *FAO1* de la presente invención y tenía 85% de identidad y 93% de similitud con los genes *FAO2a* y *FAO2b* de la presente invención. Aunque el gen *FAOT* era el más similar al gen *FAO2a* de la presente invención, la diferencia era aún equivalente a aproximadamente 49 aminoácidos de los 704. Estos datos demuestran que el gen *FAOT* publicado es ligeramente más similar a cualquiera de los genes *FAO* de la presente invención de lo que los genes *FAO* de la presente invención son entre sí. Estos datos indican además la probabilidad de que *FAO1*, *FAO2* y *FAOT* son genes diferentes, en vez de alelos entre sí.

20 Aunque se obtuvo un fragmento PCR de 1200 pb a partir de todos los clones utilizando iniciadores 30F2 y 30R1, la banda era generalmente muy débil, incluso después de haber realizado la optimización de las condiciones de la PCR. Cuando los iniciadores, que fueron diseñados a partir de la secuencia *FAOT* publicada, fueron alineados con los genes *FAO1*, *FAO2a* y *FAO2b* de la presente invención (Tabla 4), se observó una carencia significativa de homología, particularmente con *FAO-F1* y *FAO-F2*. Esto indica también que *FAO1*, *FAO2* y *FAOT* son genes diferentes. Incluso aunque se utilizó la información publicada de la secuencia de *Candida tropicalis* para obtener los iniciadores de la PCR, el grado de variación de la secuencia hizo encontrar un incierto par en funcionamiento del iniciador de la PCR.

Tabla 4. Comparación de la secuencia del iniciador con la secuencia correspondiente en *FAO1* y *FAO2* de *Candida tropicalis*.

Gen	Comparación de secuencia	Iniciador
<i>FAOT</i> de <i>C. t.</i>	352 5' <b>TCG TGG CGT GAC TCT CCT</b> 3' 369 (SEQ ID NO: 35)	FAO-F1
<i>FAO1</i> de <i>C. c.</i>	352 5' <b>TCG TGG CGT GAC TCC CCT</b> 3' 369 (SEQ ID NO: 44)	
<i>FAO2a</i> de <i>C. c.</i>	352 5' <b>TCI TGG CGT GAI TCC CCG</b> 3' 369 (SEQ ID NO: 45)	
<i>FAO2b</i> de <i>C. t.</i>	352 5' <b>GCC TGG CGT GAI TCC CCG</b> 3' 369 (SEQ ID NO: 46)	
<i>FAOT</i> de <i>C. t.</i>	608 5' <b>CTG GTG CTG GTG TAG T</b> 3' 623 (SEQ ID NO: 38)	FAO-F2
<i>FAO1</i> de <i>C. c.</i>	608 5' <b>CCG GTG CTG GTG TCG T</b> 3' 623 (SEQ ID NO: 47)	
<i>FAO2a</i> de <i>C. c.</i>	608 5' <b>CCG GTG CTG GTG TCA T</b> 3' 623 (SEQ ID NO: 48)	
<i>FAO2b</i> de <i>C. t.</i>	608 5' <b>CCG GTG CTG GTG TCA T</b> 3' 623 (SEQ ID NO: 49)	
<i>FAOT</i> de <i>C. t.</i>	1780 5' <b>TTG GTA CCC ATG CTT GTG G</b> 3' 1762 (SEQ ID NO: 41)	FAO-R1
<i>FAO1</i> de <i>C. c.</i>	1780 5' <b>TTG GCA CCC ATG GTT GGG G</b> 3' 1762 (SEQ ID NO: 50)	
<i>FAO2a</i> de <i>C. c.</i>	1780 5' <b>TTG GCA CCC ATG GCT GTG G</b> 3' 1762 (SEQ ID NO: 51)	
<i>FAO2b</i> de <i>C. t.</i>	1780 5' <b>TTG GCA CCC ATG CCT GTG G</b> 3' 1762 (SEQ ID NO: 52)	

#### EJEMPLO 4

##### 5 Subclonación y expresión de *FAO1* y *FAO2* en *E. coli*.

###### PCR de genes *FAO1* y *FAO2*

Ya que los datos de homología de secuencia indican fuertemente que *FAO1* y *FAO2* eran genes diferentes, se investigó la unicidad de los dos genes por medio de clonación y expresión de *FAO1* y *FAO2a* individualmente en *E. coli* para determinar la especificidad por el sustrato de los dos producto génicos.

10 Los marcos de lectura abiertos de estos dos genes fueron amplificados por medio de PCR y clonados (Argonne National Labs, Argonne, IL) en el vector de autorreplicación pJF 118EH (13). Este vector, que contiene ya sea al gen *FAO1* o al gen *FAO2*, fue transformado en *E. coli*. La colocación de los genes para las enzimas *FAO1* y *FAO2* en *E. coli* permitió la sobreexpresión de las proteínas permitiendo así la generación de grandes cantidades de enzimas en un fondo limpio de tal manera que sus propiedades podrían ser más claramente definidas.

15 Inicialmente, se determinaron únicamente las propiedades de la enzima derivada de la secuencia nativa de *FAO2a*. Se sabe, sin embargo que *FAO2a* tiene un codón CTG, que es traducido como una serina en *C. tropicalis* pero como una leucina en *E. coli*. Debido a esto, se generó un constructo de *FAO2a* (denominado *FAO2a'*) que tiene un codón TCG, que codifica para serina tanto en *C. tropicalis* como en *E. coli*, en lugar del codón CTG. Las propiedades tanto de *FAO2a* como de *FAO2a'* se determinaron como se describe en los siguientes ejemplos.

20 Los iniciadores utilizados para amplificar las regiones de codificación de los genes *FAO1* y *FAO2* por medio de PCR se muestran más adelante. Se añadieron los sitios de restricción (subrayados) *EcoRI* y *BamHI* en los extremos 5' y 3', respectivamente. El codón de inicio ATG en el iniciador hacia adelante y los codones de terminación duales en los iniciadores inversos se muestran en cursiva.

## ES 2 379 776 T3

FAO1U 5' -CCGAATTCGACATGGCTCCATTTTTG-3' (SEQ ID NO: 53)

FAO1L 5'-CCGGATCCATTACTACAACCTTGGCCTTGGT -3' (SEQ ID NO: 54)

FAO2U 5'-CCAGTGAATTCAGATGAATACCTTCT-3' (SEQ ID NO: 55)

FAO2L 5'-CCGGATCCCCGTCTCACTACAACCTTG-3' (SEQ ID NO: 56).

- 5 PCR utilizó Platino Pfx ADN Polimerasa de Life Technologies, Inc. (Rockville, MD 20849 - 6482). Las condiciones de reacción para cada reacción de 50 µl fueron:

amortiguador 1X (suministrado por el fabricante)

MgSO<sub>4</sub> 1,0, 1,5 ó 2,0 mM

1 µM cada iniciador -

- 10 0,2 mM cada uno de los 4 dNTP

200 - 400 ng del plásmido FAO1 o FAO2

1 unidad de Platino Pfx Polimerasa

- 15 Se incubaron las reacciones en un termociclador Robocycler Gradient 96 (Stratagene) durante un ciclo a 94°C por 2 minutos, seguido por 30 ciclos a 94°C (30 segundos); 55°C, 58°C ó 61°C (45 segundos); 72°C (2 minutos). Se completaron las reacciones por incubación a 72°C (10 minutos) durante 1 ciclo.

*FAO1* produjo el producto esperado (2,1 kb) bajo todas las condiciones ensayadas. Las condiciones óptimas para *FAO2* fueron con MgSO<sub>4</sub> 1,0 - 1,5 mM y 55°C - 58°C.

Purificación en gel de agarosa de productos PCR

- 20 Se reunieron y purificaron tres de las reacciones PCR para cada gen con el kit de Purificación por PCR QIAquick-spin (Qiagen) siguiendo las instrucciones del fabricante. Se fraccionó luego el ADN sobre un gel de agarosa al 1,0 %. Se removieron las bandas de 2,1 kb y se extrajo el ADN con el Kit de Extracción en Gel QIAEX II (Qiagen) siguiendo las instrucciones del fabricante.

Ligación dentro de pJF118EH

- 25 Se digirió el vector de expresión pJF118EH (13) con *EcoRI* y *BamHI*, y se fraccionó sobre un gel de agarosa al 1,0%. Se cortó y purificó la banda con el Kit de Extracción en Gel QIAEX II (Qiagen). Se digirieron los productos PCR *FAO1* y *FAO2* con *EcoRI* y *BamHI*, y se purificó el gel en la misma forma. Se visualizaron las bandas digeridas sobre un gel de agarosa pero no se determinaron las concentraciones exactas.

- 30 Se incubaron las reacciones de ligación (20 µl) que contenían 1 µl de *FAO1* o *FAO2* y 4 µl de pJF1 18EH en 1X amortiguador de ligación (Promega) con 1 µl (3 unidades) T4 ADN ligasa (Promega) durante 2 horas a 25°C. Se transformaron 100 µl de Library Efficiency DH5 *E. coli* (Life Technologies, Inc.) con 1,5 µl de cada reacción de ligación.

Se seleccionaron seis colonias de cada transformación, se hicieron minipreparaciones de ADN y se determinó el tamaño del inserto por medio de digestión con *EcoRI* y *BamHI*. 5 de los 6 clones *FAO1* contenían el inserto de 2,1 kb. Todos los 6 clones *FAO2* contenían al inserto de 2,1 kb.

- 35 Se seleccionó un clon de cada uno y se almacenaron patrones de glicerol a -80°C. Se utilizaron también muestras de cada clon para análisis de la expresión y actividad de la enzima. El clon *FAO1*, denominado FAO1-EC, contenía al plásmido FAO1jf (ver la Figura 1), y el clon *FAO2*, denominado FAO2-EC, contenía a plásmido FAO2jf (ver la Figura 2). Se confirmó la secuencia de ambos genes *FAO* en éstos plásmidos por parte de Sequetech Corporation, Mountain View, CA.

- 40 Inducción de *FAO1* y *FAO2*

Los cultivos durante la noche de FAO1-EC y FAO2-EC se desarrollaron a 30 °C en 5 ml de Caldo Terrific (TB)

(Sigma Chemical Co, St. Louis, MO) más 100 µg/ml de ampicilina a 250 rpm. Se colocaron 50 ml de TB más 100 µg/ml de ampicilina en dos matraces deflectores de 500 ml. Se utilizó TB más 100 µg/ml de ampicilina tanto para los cultivos iniciadores como los cultivos que fueron inducidos para producir la enzima. Se inocularon los matraces con los cultivos de la noche con una absorbancia a 600 nm de aproximadamente 0,2. Los cultivos se desarrollaron a 30 °C, con agitación a 250 rpm. Cuando cada cultivo había alcanzado una absorbancia a 600 nm de 5 a 6, fue inducido con IPTG hasta una concentración final en cada cultivo de 1 mM. Se permitió luego que los fueran incubados otras tres horas después de la inducción. Se recolectaron las células por centrifugación (Sorvall RC5C) aproximadamente a 6.000 x g (rotor GS3 a 6.000 rpm) durante 10 minutos. El caldo sobrenadante agotado (SB) fue removido y recuperado, y se congelaron los precipitados celulares a -20°C para uso posterior.

10 Preparación de extractos celulares de *E. coli*

Se prepararon microsomas para ensayos de alcohol oxidasa resuspendiendo los precipitados celulares inducidos en 50 ml de amortiguador de PO<sub>4</sub>/glicerol. Se añadieron 500 µl de una solución PMSF 100 mM para una concentración final de 1 mM de PMSF en la suspensión. Se rompieron las células pasándolas dos veces a través de una celda francesa presurizada refrigerada, y fueron examinadas por medio de un microscopio de contraste de fases para garantizar que 95% o más de las células habían sido rotas. Se centrifugó la suspensión de células rotas (BCS) aproximadamente a 37.000 x g durante 30 minutos para precipitar los residuos celulares. Se removió el sobrenadante (extracto libre de células, CFE) y guardó para los análisis. El precipitado de residuos celulares (CD) fue resuspendido en 50 ml de amortiguador de fosfato-glicerol para que fuera igual a la concentración original de las células en el cultivo.

20 Preparación de suspensiones microsomales de *E. coli*

Se calentaron los extractos libres de células hasta temperatura ambiente y se mezclaron completamente antes de remover 3 ml de muestra de cada uno para la preparación microsomal. Estas muestras fueron centrifugadas en una ultracentrífuga a 100.000 x g durante una hora a 4°C para precipitar los microsomas. El sobrenadante fue luego removido y recuperado, y se resuspendió cada precipitado microsomal en 0,5 ml del amortiguador de PO<sub>4</sub>/glicerol para una concentración 6x de la enzima FAO.

**EJEMPLO 5**

Construcción de un gen *FAO2a* alterado en el codón CTG

Se llevó a cabo una alteración de un codón de *FAO2a* por medio de PCR de extensión por superposición. Se utilizaron los siguientes grupos de iniciadores para generar los constructos iniciales:

30 A9.1N 5' ATC AAC GCC ACC CCA ACC 3' (SEQ ID NO: 57)

FAO2-CTG-R 5' GGT TTC TCC ATA AAC GAG TAC CTG AAAGGG TCA ACC 3' (SEQ ID NO: 58)

Estos iniciadores fueron diseñados para cubrir la región 3' de 208 pb hasta 3' de 545 pb del inicio del ORF produciendo un fragmento de 338 pb de longitud. La alteración del codón CTG por CGA (TCG en el complemento inverso) se indica en negrilla. Un sitio de restricción anterior *KpnI* (GGTACC), anteriormente subrayado, ha sido eliminado por medio de esta alteración del codón.

La PCR utilizó Platino Pfx ADN Polimerasa de Life Technologies, Inc. Las condiciones de reacción para cada reacción de 50 µl eran:

amortiguador 1X (suministrado por el fabricante)

MgSO<sub>4</sub> 1,0 mM

40 1 µM cada iniciador

0,3 mM cada uno de los 4 dNTP

200 - 400 ng del plásmido *FAO2a*

1 unidad de Platino Pfx Polimerasa

A9.1E 5' ATC TGT CTA GCA AAG GTC 3' (SEQ ID NO: 59)

## ES 2 379 776 T3

FAO2-CTG-F 5' GGT TGA CCC TTT CAG GTA CTC GTT TAT GGA GAA ACC 3' (SEQ ID NO: 60)

Estos iniciadores fueron diseñados para cubrir la región 3' de 510 pb hasta 3' de 2069 pb del inicio del ORF produciendo un fragmento de 1560 pb de longitud. Los FAO2-CTG-F y FAO2-CTG-R son iniciadores de superposición.

- 5 Este PCR también utilizó Platino Pfx ADN Polimerasa de Life Technologies, Inc. Las condiciones de reacción para cada reacción de 50 µl eran:

amortiguador 1X (suministrado por el fabricante)

MgSO<sub>4</sub> 5,0 mM

1 µM cada iniciador

- 10 0,3 mM cada uno de los 4 dNTP

200 - 400 ng del plásmido FAO2a

1 unidad de Platino Pfx Polimerasa

- 15 Las reacciones fueron incubadas en un termociclador PTC200 (MJ Research) para un ciclo a 94°C durante 2 minutos, seguido por 30 ciclos a 94°C (15 segundos); 55°C (30 segundos); 68°C (4 minutos). Las reacciones fueron completadas por incubación a 72°C (7 minutos) durante 1 ciclo.

Los fragmentos PCR de tamaño apropiado fueron separados sobre gel de agarosa de bajo punto de fusión al 1%, y fueron luego cortados en tubos de microfuga separados. Se realizó una tercera PCR utilizando iniciadores A9.1N y A9.1E. Esta PCR también utilizó Platino Pfx ADN polimerasa de Life Technologies, Inc. Las condiciones de reacción para cada reacción de 50 µl eran:

- 20 amortiguador 1X (suministrado por el fabricante)

MgSO<sub>4</sub> 2,0 mM

1 µM cada iniciador

0,3 mM cada uno de los 4 dNTP

1 µl de cada fragmento PCR (338 pb y 1560 pb, agarosa de bajo punto de fusión fundida a 65 °C)

- 25 1 unidad de Platino Pfx Polimerasa

- 30 Se incubó la reacción en un termociclador PTC200 (MJ Research) para un ciclo a 94°C durante 2 minutos, seguido por 30 ciclos a 94°C (15 segundos); 60°C (30 segundos); 68°C (2 minutos). Las reacciones fueron completadas por incubación a 72°C (7 minutos) durante 1 ciclo. Esto dio como resultado un fragmento PCR que tenía 1862 pb de longitud y cubierto de 208 pb hasta 2069 pb desde el inicio del ORF. Este fragmento era TOPO-TA clonado y el plásmido resultante fue preparado como se describe en Materiales y Métodos (Ejemplo 1).

- 35 Con el propósito de reemplazar el codón CTG en el plásmido FAO2jf con la secuencia modificada, se cortó el plásmido, FAO2jf, con *KpnI* y *MfeI* para remover un fragmento de 470 pb de longitud, dejando la porción mayor del plásmido, 6924 pb intacta. Se cortó también el fragmento de 1862 pb con *KpnI* y *MfeI* para remover un fragmento de 470 pb de longitud. Se purificaron los fragmentos apropiados, el fragmento de 6924 pb del plásmido FAO2j f y el fragmento de 470 pb cortado del fragmento de 1862 pb que contenía el codón CTG modificado, sobre gel de agarosa al 1%, y luego fueron extraídos del gel utilizando el kit Qiaquick de Qiagen.

- 40 Después de obtener los fragmentos apropiados de ADN, se estableció una ligación utilizando el producto y el protocolo de Ligación Quick de New England Biolabs. Esta ligación fue transformada nuevamente en células DH5α. Se seleccionaron preparaciones de plásmido de clones putativos cortando con *KpnI* y *MfeI*. Se seleccionó uno de los clones positivos para estudios adicionales y fue denominado como FAO2a'EC. Ya que todo el gen *FAO2α* modificado en CTG (denominado *FAO2α* excepto por el fragmento de 469 pb que contenía la modificación del codón cuya secuencia había sido previamente modificada) únicamente se verificó la secuencia de esta porción de 469 pb del gen. El gen tenía el codón apropiado cambiado de un CTG (pb 2049 - 2051 en la SEQ ID NO: 9).

También tenía una segunda mutación, presumiblemente generada por PCR. La mutación era en pb 2117 y dio como resultado un cambio de codón de ATT a ATC, ambas codifican para isoleucina.

Preparación del casete de amplificación para *FAO1* y *FAO2*

- 5 La etapa inicial en la creación del casete de amplificación de *FAO1* y *FAO2* fue para utilizar iniciadores con sitios de restricción para *PacI* sobre los extremos para amplificar el ORF y las regiones secuencia arriba de cada gen. Para *FAO1* los iniciadores utilizados fueron:

(sitio de restricción para *PacI* subrayado)

FAO1-F3-*PacI* 5' CCT TAA TTA ATG CAT ACT CGG AGC ATA TCG C 3' (SEQ ID NO: 61)

FAO1-R3-*PacI* 5' CCT TAA TTA ATG GGC GGA ATC AAG TGC C 3' (SEQ ID NO: 62)

- 10 La región cubierta es la 5' de 1939 pb y la 3' de 245 pb del ORF

Para *FAO2* los iniciadores fueron:

FAO2-F1-*PacI* 5' CCT TAA TTA ATC TCA CCA AGT ACG AGA ACG 3' (SEQ ID NO: 63)

FAO2-R1-*PacI* 5' CCT TAA TTA AGA CGC AAG CAC AGG TGC C 3' (SEQ ID NO: 64)

La región cubierta es la 5' de 1520 pb y la 3' de 526 pb del ORF

- 15 Se utilizó el kit Accutaq LA Core (Sigma, #ACCUCORE) para crear los fragmentos PCR. Este kit contiene tanto una Taq ADN polimerasa con una pequeña cantidad de una enzima de corrección de prueba. La mezcla de reacción fue preparada siguiendo las recomendaciones en el kit. Las condiciones seguidas para la PCR fueron: 98°C durante 30 s seguido por 35 ciclos de: 94°C durante 5 s, 65°C durante 20 s y 68°C durante 5 min. Después de esto hubo una prolongación final de 72°C durante 5 min. Estos fragmentos PCR fueron clonados en TA-TOPO dentro del vector PCR 2.1, que fue transformado en células Top10F' de la cepa de *E. coli*, creando las cadenas *FAO1:Pacl* y *FAO2:Pacl*.
- 20

- 25 Con el propósito de reemplazar la mayor parte de la secuencia del ORF obtenida por medio de PCR con ADN genómico, se cortó *FAO1:Pacl* con *AatII* y *ZheI* para liberar un fragmento aproximadamente de 1,9 kb de longitud. Su clon de la biblioteca complementaria, A8.1, también fue cortado con *AatII* y *NheI*. *FAO2:Pacl* y su clon de la biblioteca complementaria, A9.1, fueron cortados con *NheI* y *BstBI* para liberar un fragmento de 3,2 kb. Los fragmentos apropiados fueron purificados sobre gel de agarosa de bajo punto de fusión al 1%, y luego fueron extraídos del gel utilizando el kit Qiaquick de Qiagen.

- 30 Después de obtener los fragmentos apropiados de ADN, se estableció una ligación tanto para *FAO1* como para *FAO2* en la cual el ORF de los clones de la biblioteca estaba ligado dentro del vector PCR 2.1 con las regiones secuencia arriba y secuencia abajo del gen. Estas ligaciones fueron transformadas nuevamente dentro de células Top 10F. Se seleccionaron las preparaciones de plásmido a partir de clones putativos cortando con *AatII* y *NheI* para los clones *FAO1*, o cortando con *NheI* y *BstBI* para los clones *FAO2*, y cortando con *PacI* para todos los clones.

- 35 Debido a que el comienzo y el final de cada ORF fue obtenido a partir de la amplificación original por PCR de cada gen, estas áreas deben ser secuenciadas para estar seguros de que no se habían introducido mutaciones dentro del ORF. Se enviaron tres clones de reemplazo del ORF *FAO1:Pacl* a Sequetech Corp. (Mountain View CA) para secuenciación. Se enviaron también cinco clones de reemplazo del ORF *FAO2:Pacl* a Sequetech Corp. para secuenciación. Los clones de reemplazo de ORF *FAO1: Pacl* 1 y 2, y los clones de reemplazo del ORF *FAO2:Pacl* 1 - 5 eran todas libres de mutaciones, y podrían ser utilizadas para la elaboración de casetes de amplificación.

## EJEMPLO 6

- 40 Preparación de un constructor de fusión del promotor

Las fusiones entre el promotor del gen *CYP52A2* para citocromo p450 monooxigenasa, y el ORF ya sea de *FAO1* o de *FAO2* fueron preparadas por medio de PCR de prolongación de la superposición. Se establecieron las reacciones PCR tanto para el promotor *CYP52A2* como para el ORF de *FAO* utilizando el kit de PCR Accutaq. Los iniciadores utilizados para obtener el promotor *CYP52A2* para la fusión con *FAO1* fueron (sitio de restricción de *PacI*):



00218-179A 5' CCT TAA TTA AAG TCT CCA AGT TGA CCG AC 3' (SEQ ID NO: 65)

FAO1-A2-R 5' AAA TGG AGC CAT GGT CGT GAT GTG TG 3' (SEQ ID NO: 66)

5 Se diseñó *FAO-A2-R* para montarlo en el extremo 3' del promotor *CYP52A2*, pero la última mitad del extremo 5' a 3' del iniciador se derivó de la secuencia complementaria del inicio del ORF de *FAO1*. Contrariamente, los iniciadores utilizados para amplificar el ORF de *FAO1* fueron:

FAO1-A2-F 5' CAC ATC ACG ACC ATG GCT CCA TTT TTG CC 3' (SEQ ID NO: 67)

FAO1-R3-Pac I 5' CCT TAA TTA ATG GGC GGA ATC AAG TGC C 3' (SEQ ID NO: 68)

10 Se diseñó *FAO1-A2-F* para montarlo en el extremo 5' del ORF de *FAO1*, con la primera mitad del extremo 5' a 3' del iniciador derivada de la secuencia complementaria del extremo del promotor *CYP52A2*. De este modo, el promotor *CYP52A2* y las regiones producidas del ORF de *FAO* tenían secuencias complementarias entre sí.

Los iniciadores utilizados para obtener el promotor *CYP52A2* fueron:

00218-179A 5' CCT TAA TTA AAG TCT CCA AGT TGA CCG AC 3' (SEQ ID NO: 69)

FAO2-A2-R 5' GAA GGT ATT CAT GGT CGT GAT GTG TG 3' (SEQ ID NO: 70)

Los iniciadores utilizados para las regiones del ORF de *FAO2* fueron:

15 FAO2-A2-F 5' CAC ATC ACG ACC ATG AAT ACC TTC TTG CC 3' (SEQ ID NO: 71)

FAO2-R1-Pac I 5' CCT TAA TTA AGA CGC AAG CAC AGG TGC C 3' (SEQ ID NO: 72)

20 Una vez se habían obtenido las bandas PCR apropiadas, se purificó cada fragmento PCR sobre un gel de agarosa de bajo punto de fusión al 1%, que luego fue cortado del gel, y se extrajo el producto PCR. La banda del promotor *CYP52A2* era aproximadamente de 1,1 kb para ambas reacciones de fusión. La banda del ORF para *FAO1* era aproximadamente de 2,4 kb y la banda del ORF para *FAO2* era aproximadamente de 2,7 kb. Se estableció la PCR de fusión utilizando una pequeña porción del promotor *CYP52A2* y las bandas purificadas en gel del ORF de *FAO* como ADN molde. Se utilizó el kit Accutaq, y los iniciadores fueron el iniciador hacia adelante para el promotor *CYP52A2* y el iniciador inverso para cada ORF de *FAO*. Durante la reacción PCR, el promotor *CYP52A2* y el ORF de *FAO* hibridaron juntos en sus extremos complementarios, y la reacción PCR continuó sobre la unión para formar el constructo de fusión. La banda de la fusión *pCYP52A2FAO1* era de 3,5 kb y la banda de la fusión *pCYP52A2FAO2a* era de 3,8 kb.

25 Estas bandas de fusión fueron luego clonadas TA-TOPO y transformadas en células de *E. coli* Top10F'. Nuevamente, la mayor parte del marco ORF de cada clon de fusión fue reemplazado con el correspondiente ADN genómico debido a la posibilidad para introducir mutaciones de la reacción PCR dentro del ORF. Los clones *pCYP52A2FAO1* fueron cortados con las enzimas de restricción *AatII* y *NheI* para remover la mayor parte del ORF de *FAO1*. Los clones *pCYP52A2FAO2a* fueron cortados con las enzimas de restricción *AatII* y *BstBI*. Se obtuvo un fragmento genómico correspondiente para cada ORF de *FAO* en una forma similar. Después de la digestión, se purificaron los fragmentos apropiados de cada gen y se los ligo entre sí. El plásmido resultante contenía el fragmento *pCYP52A2FAO* con el ADN genómico de *FAO* reemplazado en el vector PCR 2.1. Este fue transformado dentro de células Top10F" y tres de los clones *pCYP52A2FAO1* y cinco de los clones *pCYP52A2FAO2a* fueron enviados a Sequetech Corporation para secuenciación.

30 Después de la secuenciación, se demostró que los tres clones de reemplazo del ORF de *pCYP52A2FAO1* no tenían mutaciones y eran adecuados para elaborar casetes de amplificación. La secuenciación demostró que todos los cinco clones de reemplazo del ORF de *pCYP52A2FAO2a* tenían mutaciones aparentes introducidas por la PCR en diferentes lugares en el ORF.

35 La creación de casetes de amplificación para el clon de fusión *pCYP52A2FAO1* fue lograda por medio de la remoción del fragmento del vector PCR 2.1 y clonándolo dentro del vector pURA3in. Este vector consistía de 193 pNEB con fragmentos invertidos del gen URA3 de *C. tropicalis* que flanquea cada región del sitio de inserción. El sitio de inserción es un sitio único para la enzima de restricción *PacI*, dentro de la cual se puede insertar el clon de fusión *pCYP52A2FAO1* después de la digestión de restricción con *PacI* y la purificación en gel. Los fragmentos del gen URA3 están flanqueados por los sitios de digestión de restricción de *AsclI* y *PmeI* de tal manera que el fragmento completo de inserción del clon *FAO* flanqueado por las secuencias del gen invertido URA3 puede ser removido y utilizado para transformar una cepa base *Candida tropicalis ura-*. Esto fue exitosamente realizado utilizando *pCYP52A2FAO1*.

**EJEMPLO 7**

## Resultados

## Actividad de la alcohol oxidasa en muestras de fermentación

5 Durante el transcurso de una fermentación diácida con la cepa HDC23-3 utilizando HOSFFA como sustrato, la actividad de FAO generalmente no apareció hasta aproximadamente 2 - 4 horas después de la inducción con HOSFFA (ver la Figura 3). La actividad aumentó rápidamente hasta alcanzar un pico aproximadamente 30 - 40 horas después de la inducción, y luego cayó aproximadamente 5 a 10 veces por 100 - 120 horas después de la inducción. Aunque esta tendencia se mantuvo constante a medida que avanzaba la fermentación, el tiempo y la altura del pico de actividad varió algo.

10 Una de las dificultades encontradas en la medición de la actividad de FAO en los extractos fue la presencia de catalasa en las células. La catalasa convierte peróxido de hidrógeno en la célula en agua y oxígeno molecular. La catalasa en los extractos podría competir con la peroxidasa de rábano por el peróxido de hidrógeno producido por la oxidación del alcohol hasta el aldehído por la FAO, lo cual daría como resultado mediciones de actividad que fueran menores que la actividad real. No se encontró actividad de FAO en el sobrenadante microsomal, pero hubo actividad  
15 en el precipitado microsomal resuspendido, y esta actividad fue aproximadamente la misma que aquella en el extracto original. La actividad de la catalasa estaba presente en el precipitado microsomal, pero aproximadamente en 1/100 del nivel encontrado en el extracto original. Estos datos indican que, bajo las condiciones de ensayo utilizadas, la catalasa no compitió significativamente con la peroxidasa de rábano por el peróxido de hidrógeno en la mezcla de reacción.

**EJEMPLO 8**

## Fraccionamiento de la actividad del alcohol oxidasa

25 Las células de *E. coli* que contenían los plásmidos, FAO1jf o FAO2jf, fueron inducidas para expresar FAO1 o FAO2 como se describe en Materiales y métodos (Ejemplo 1). El caldo agotado, el extracto libre de células, y la resuspensión de los residuos celulares (ver Materiales y métodos, Ejemplo 1) fueron todos analizados por la actividad de FAO en un esfuerzo para determinar la ubicación de la actividad de FAO. Si se deposita en forma extracelular, la actividad de FAO sería observada en el caldo agotado. Si se deposita internamente, la actividad de FAO sería detectada en el extracto libre de células o en los residuos celulares resuspendidos. La enzima soluble o la enzima microsomal enlazada a la membrana se encontrarían en el extracto libre de células. Para determinar la ubicación en la célula de la alcohol oxidasa, se llevaron a cabo ensayos de actividad con las diversas fracciones. No se observó actividad de FAO en el caldo agotado, y la actividad encontrada en la resuspensión del precipitado de  
30 residuos celulares (FAO1 = 0,002 U/ml; FAO2 = 0,002 U/ml) fue de una quinta parte hasta una décima parte de la actividad determinada para el extracto libre de células (FAO1 = 0,026 U/ml; FAO2 = 0,044 U/ml).

35 Las preparaciones microsomales fueron elaboradas a partir de los extractos libres de células de estos cultivos de *E. coli* de FAO1 y FAO2. La actividad en cada precipitado microsomal resuspendido y sobrenadante fue analizada por la actividad de FAO. No se observó actividad en el sobrenadante del precipitado microsomal de FAO1, y la actividad en el sobrenadante del precipitado microsomal de FAO2 (0,010 U/ml) fue aproximadamente una quinta parte de la actividad en el precipitado microsomal de FAO2 resuspendido (0,050 U/ml). La actividad en la preparación microsomal de FAO2 fue casi dos veces la actividad que fue observada en la preparación microsomal de FAO1 (0,031 U/ml). Estos datos indican que la mayor parte de la actividad enzimática estaba contenida en los microsomas aislados a partir de los cultivos de *E. coli*.

## Optimización de la inducción

45 Con el propósito de determinar la temperatura óptima para la síntesis de las enzimas FAO1 y FAO2a en cultivos de *E. coli*, se hicieron cultivos de semilla durante la noche a 37 °C y 250 rpm en caldo Terrific. Se añadieron 50 ml de caldo Terrific a cada uno de los cuatro matraces deflectores de 500 ml; uno para cada uno de FAO1 y FAO2a que son incubados ya sea a 30 °C ó 37 °C. Cada matraz fue inoculado con los cultivos realizados durante la noche con una absorbancia a 600 nm de entre 0,7 y 0,8, y luego fue incubado con agitación a 250 rpm. Cuando cada cultivo había alcanzado u a absorbancia a 600 nm de aproximadamente 4,5, fue inducido con IPTG hasta una concentración final de 1 mM. Se recolectaron las células 3 horas después de la inducción por medio de centrifugación a 6000 x g durante 10 minutos. Se removió el sobrenadante y se almacenaron las células a -20 °C.  
50 Los cultivos de *E. coli* de FAO1 y FAO2a que se desarrollaron a 30 °C y 37 °C fueron preparados para ensayos de la enzima como se describe en Materiales y métodos. El extracto libre de células preparado a partir de cada uno de estos cultivos fue analizado por la actividad de FAO. Se encontró que el cultivado a 30 °C dio como resultado una mayor producción de FAO1 (14,7 U/ml) y de FAO2 (2,44 U/ml) que el cultivado a 37 °C (8,8 U/ml para FAO1 y 1,33 U/ml para FAO2a).

5 En un segundo grupo de experimentos, se determinó la concentración óptima del inductor, IPTG, para *E. coli* que contenía ya sea *FAO1* o *FAO2a*. El experimento de crecimiento fue creado de la misma manera que el experimento anterior, con cultivos de semilla durante la noche de *FAO1* y *FAO2a* que se desarrollaron a 37°C en Caldo Terrific más 100 µg/ml de ampicilina. Se añadieron 50 ml de caldo Terrific a cada uno de los seis matraces; tres matraces cada uno para *FAO1* y *FAO2a*. Un matraz de cada grupo fue inducido ya sea con IPTG 0,5 mM, IPTG 1,0 mM, o IPTG 2,0 mM. Los cultivos se desarrollaron a 30°C con agitación a 250 rpm. Cuando cada cultivo había alcanzado una absorbancia a 600 nm de aproximadamente 5,0, fue inducido con la concentración apropiada de IPTG. Se permitió que los cultivos se desarrollaran durante otras 2,5 horas después de la inducción antes de recolectar las células por medio de centrifugación a 6000 x g durante 10 minutos. Se almacenaron los precipitados celulares a -20°C. Se prepararon extractos libres de células y después de analizar la actividad de la alcohol oxidasa, se encontró que existía poca dependencia de la actividad de *FAO1* sobre la concentración de IPTG (0,5 mM = 13,24 U/ml; 1 mM = 13,89 U/ml; 2 mM = 14,3 U/ml). Una respuesta similar a la concentración de IPTG fue observada para *FAO2a* (0,5 mM = 4,74 U/ml; 1 mM = 5,68 U/ml; 2 mM = 4,65 U/ml). Se escogió una concentración estándar de IPTG de 1 mM para la inducción de *FAO1*, *FAO2a* y *FAO2a'*.

15 Tabla 5. Alcoholes utilizados para analizar la especificidad por el sustrato

Actividad Detectada*	Compuesto Analizado	Actividad Detectada	Compuesto Analizado
	2-Pentanol		1-Fenilpropan-1-ol
	2-Hexanol		3-Fenilpropan-1-ol
X	2-Decanol		2-Fenilbutan-1-ol
X	2-Undecanol		Metanol
X	2-Dodecanol		Etanol
X	2-Hexadecanol		Propanol
	3-Octanol		Butanol
X	10-Undecen-1-ol	X	Hexanol
	1,8-Octanodiol	X	Octanol
X	1,2-Octanodiol	X	Decanol
X	1,10-Decanodiol	X	Dodecanol
X	1,2-Dodecanodiol	X	Tetradecanol
X	1,16-Hexadecanodiol	X	Hexadecanol
X	Ácido 10-Hidroxicanoico		4-ciclohexil-1-butanol
X	Ácido 12-Hidroxicanoico		3-ciclohexil-1-propanol
X	Ácido 16-Hidroxicanoico		2-ciclohexiletanol
	Citronelol		Ciclohexilmetanol
	Geraniol		
	Linalool		

\* "X" = Actividad detectada utilizando este sustrato ya sea con *FAO1* o con *FAO2*

## Especificidad del sustrato de FAO1, FAO2a y FAO2a'

5 FAO1, FAO2a y FAO2a' fueron analizados por su actividad con diversos alcoholes. Los alcoholes mostrados en la Tabla 5 fueron preparados en soluciones patrón 20 mM en acetona. Se indican los alcoholes que muestran actividad ya sea con FAO1 o con FAO2a. La concentración final del alcohol utilizado en el experimento de especificidad por el sustrato era 200  $\mu$ M. Estos mismos alcoholes fueron utilizados para determinar la  $K_m$  y la  $V_{m\acute{a}x}$  de FAO1, FAO2a, y FAO2a'. Se reportaron los perfiles de especificidad por el sustrato de FAO1, FAO2a, y FAO2a' como porcentajes de actividad para dodecanol, con dodecanol arbitrariamente fijado en una actividad del 100%. La actividad de FAO1, FAO2a, y FAO2a' con 1-alcanoles es mostrada en la Figura 4. En forma muy interesante, FAO1 prefirió 1-octanol como sustrato, siendo 1-tetradecanol el alcohol preferido de cadena más larga. En contraste, FAO2a y FAO2a' prefirieron 1-dodecanol por encima de todos los demás 1-alcanoles. Hubo una gran caída en la actividad al pasar de un alcohol C8 a un alcohol C6 con FAO1, FAO2a, y FAO2a'. La actividad de las tres enzimas sobre 2-alcanoles es mostrada en la Figura 5. FAO2a y FAO2a' oxidan 2-octanoles bastante bien; sin embargo no se detectó actividad con FAO1. En contraste, FAO1 oxida bien ácidos grasos  $\omega$ -hidroxi, pero no se midió actividad para ácidos grasos  $\omega$ -hidroxi con FAO2a y FAO2a' (Figura 6). Estos resultados indican que FAO1 y FAO2a parecen ser enzimas muy diferentes, con diferencias significativas en la especificidad por el sustrato. En forma muy interesante, FAO2a y FAO2a' oxidan 10-undecen-1-ol mucho mejor que FAO1. Dentro del error experimental, la especificidad por el sustrato de FAO2a y FAO2a' es esencialmente la misma, indicando que tener una serina o una leucina en la posición del aminoácido 177 tiene poco o ningún efecto sobre la especificidad por el sustrato de la enzima. Aunque la especificidad por el sustrato de FAO2b no se llevó a cabo, debido a la cercana homología con FAO2a, es probable que sea muy similar.

Los extractos elaborados a partir de muestras del fermentador inducidas con HOSFFA mostraron buena actividad con ácido 16-hidroxihexadecanoico, pero ninguna actividad con 2-dodecanol (datos no mostrados). Por lo tanto, parece que FAO1 es inducida en un grado mucho mayor que FAO2a y, el menos en fermentaciones con HOSFFA, FAO1 parece ser la oxidasa de alcohol graso predominante.

25 Determinaciones de  $K_m$ 

Los valores de  $K_m$  indican la afinidad de una enzima por el sustrato que está siendo investigado. Los valores de  $K_m$  para la mayoría de los alcoholes que demostraron actividad con FAO1, FAO2a, y FAO2a' se determinaron por medio de la medición de la actividad de la enzima mientras se mantenía constante la concentración de la enzima y variando la concentración del alcohol añadido a la mezcla de reacción.

30 La mayor parte de los valores de  $K_m$  se determinaron a pH 7,6 utilizando el amortiguador HEPES/Triton X-100 y se disolvieron soluciones patrón de alcoholes, que fueron preparadas en concentraciones de 5 mM y/o 1 mM, en este mismo amortiguador. Para FAO2a', se determinaron los valores de  $K_m$  en HEPES 0,1 M, pH 7,6, y las soluciones patrón del sustrato se disolvieron en acetona. En este caso, todas las reacciones contenían la misma cantidad de acetona (20  $\mu$ l/ml). Los resultados se muestran en la Tabla 6.

Tabla 6. Valores de  $K_m$  para FAO1, FAO2a, y FAO2a'.

Sustrato	FAO1			FAO2a			FAO2a'		
	$K_m$	Coefficiente de Correlación	Rango de Valores Bajo Alto	$K_m$	Coefficiente de correlación	Rango de Valores Bajo Alto	$K_m$	Coefficiente de correlación	Rango de Valores Bajo Alto
Octanol	41,0 $\mu\text{M}$	0,9996	35,0 49,3	503 $\mu\text{M}$	0,9977	329,3 1069,2	654 $\mu\text{M}$	0,9976	323,5 30607,9
Decanol	15,1 $\mu\text{M}$	0,9912	11,8 21,3	74 $\mu\text{M}$	0,9990	64,3 85,9	69,6 $\mu\text{M}$	0,9978	65,1 74,8
Dodecanol	14,9 $\mu\text{M}$	0,9938	11,9 19,9	2,5 $\mu\text{M}$	0,9907	2,4 2,5	4,4 $\mu\text{M}$	0,9928	3,9 4,4
Tetradecanol	28,1 $\mu\text{M}$	0,9976	20,8 43,3	4,6 $\mu\text{M}$	0,9968	4,5 4,8	2,4 $\mu\text{M}$	0,9751	2,2 2,5
Hexadecanol	11,9 $\mu\text{M}$	0,9998	11,1 11,8	56 $\mu\text{M}$	0,9977	51,2 61,7	85 $\mu\text{M}$	0,9987	78,4 93,5
2-Decanol	NR			934 $\mu\text{M}$	0,9991	823,2 1078,5	1090 $\mu\text{M}$	0,9834	567,3 13696,9
2-Undecanol	NR			141 $\mu\text{M}$	0,9970	125,9 159,0	162 $\mu\text{M}$	0,9913	150,9 174,5
2-Dodecanol	NR			41,3 $\mu\text{M}$	0,9990	39,60 43,06	75,3 $\mu\text{M}$	0,9980	67,9 84,5
2-Hexadecanol	NR			350 $\mu\text{M}$	0,9992	332,5 370,0	720 $\mu\text{M}$	0,9982	628,1 844,2
1,2-dodecanodiol	NR			998 $\mu\text{M}$	0,9935	720,2 1622,7	928 $\mu\text{M}$	0,9969	742,6 1237,3
1,10-decanodiol	67,0 $\mu\text{M}$	0,9997	58,9 77,6	425 $\mu\text{M}$	0,9985	394,4 462,3	1607 $\mu\text{M}$	0,9975	1377,2 1927,5
1,16-hexadecanodiol	10,3 $\mu\text{M}$	0,9995	9,3 11,6	37,3 $\mu\text{M}$	0,9969	33,9 41,5	16,3 $\mu\text{M}$	0,9978	15,7 16,9
10-undecen-1-ol	9,9 $\mu\text{M}$	0,9963	8,8 11,3	45 $\mu\text{M}$	0,9986	40,7 50,1	36,1 $\mu\text{M}$	0,9916	32,7 40,4
12-hidroxi-dodecanoico	99,0 $\mu\text{M}$	0,9930	78,2 134,4	NR			NR		
16-hidroxi-hexadecanoico	7,4 $\mu\text{M}$	0,9912	5,9 9,9	NR			NR		

Para la mayoría de los alcoholes analizados, FAO1 produjo valores de  $K_m$  entre 10 - 20  $\mu\text{M}$ . El valor más bajo de  $K_m$  se encontró con el ácido 16-hidroxi-hexadecanoico. Utilizando ya sea FAO2a o FAO2a' como enzima, los valores de  $K_m$  para los diversos alcoholes fueron muy similares, demostrando nuevamente que tener una serina o una leucina en la posición del aminoácido 177 no tiene mucho efecto sobre la afinidad de la enzima por un sustrato particular.

5 Para FAO2, los valores de  $K_m$  fueron generalmente mucho más variables y produjeron valores en el rango de 2  $\mu\text{M}$  hasta superiores a 1 mM. Los sustratos con los valores más bajos de  $K_m$  fueron 1-dodecanol y 1-tetradecanol. Con base en estos resultados, es lógico que FAO1 parezca ser la oxidasa de alcohol graso predominante producida en fermentaciones de HOSFFA, ya que los ácidos grasos  $\omega$ -hidroxi son intermediarios en la oxidación de ácidos grasos hasta diácidos.

## 10 EJEMPLO 9

Amplificación de *FAO1* en *Candida tropicalis*

En fermentaciones con HOSFFA como sustrato, los ácidos grasos  $\omega$ -hidroxi son producidos en niveles entre 0,1% hasta 0,5% (p/p) en caldo de fermentación. Esto indica que existe un menor cuello de botella en la segunda etapa en la producción de diácido, es decir la conversión del alcohol hasta el aldehído. Se encontró que este ácido graso  $\omega$ -hidroxi interfiere con la purificación del diácido y causa una pérdida significativa en la recuperación. Por lo tanto, es importante reducir el nivel de ácidos grasos  $\omega$ -hidroxi producidos en la fermentación. Ya que los análisis de especificidad por el sustrato y las determinaciones de  $K_m$  realizadas con FAO1 y FAO2 indicaron que FAO1 era la oxidasa de alcohol graso predominante producida durante la fermentación con HOSFFA, se persiguió inicialmente la amplificación del gen *FAO1*.

20 Se prepararon casetes de amplificación con sitios de restricción para *PacI* en los extremos de los genes *FAO1* y *FAO2* como se describe en Materiales y métodos (Ejemplo 1). Sin embargo, ninguno de estos constructos ha sido probado en *Candida tropicalis*. Un constructo en el cual el promotor del gen *CYP52A2* reemplazó al promotor nativo en *FAO1* (SEQ ID NO: 7) fue exitosamente preparado y transformado dentro de *Candida tropicalis*. El gen *CYP52A2* codifica una citocromo P450 monooxigenasa que hace parte del complejo de hidroxilasa responsable por catalizar la primera etapa en la producción de diácido, es decir la conversión del ácido graso en el alcohol. Por medio de la fusión de la región promotora de este gen con el ORF del gen *FAO1*, se planteó la hipótesis de que el gen de la alcohol oxidasa y la correspondiente enzima serían inducidas primero y más fuertemente que podría ser de otra manera. Esto haría más rápida la conversión al diácido y ayudaría a reducir el cuello de botella en este punto de la reacción. Se seleccionaron tres clones, llamados HDC40-1, HDC40-5 y HDC40-7 para analizar en los fermentadores. Por medio de la comparación de las intensidades de las bandas en un análisis de transferencias tipo Southern y estimando los números de copias, HDC40-1 parece tener un número bajo de copias (una copia adicional de *FAO1*); HDC40-5 tiene un número mayor de copias (dos copias adicionales); y HDC40-7 tiene múltiples copias adicionales por célula.

35 Estas cepas fueron analizadas en fermentaciones con HOSFFA como sustrato. La Figura 7 muestra que, bajo condiciones de fermentación similares, todas las cepas produjeron más diácido que la cepa base, H5343. Tanto HDC40-1 como HDC40-7 tenían valores iniciales similares de productividad hasta aproximadamente 24 h del período de bioconversión. HDC40-1, sin embargo, mantuvo un nivel más alto de productividad que HDC40-7 durante las siguientes 48 h. La Figura 8 muestra el nivel de ácidos grasos  $\omega$ -hidroxi producidos en cepas HDC40, con relación a la cepa base, H5343, que no tenía genes amplificados. Obsérvese que, aunque aún existen algunos ácidos grasos  $\omega$ -hidroxi producidos, los niveles son considerablemente reducidos, comparado con H5343. Un perfil de actividad de la alcohol oxidasa (Figura 9) determinado durante el transcurso de la fermentación demuestra que la alcohol oxidasa, como se esperaba, había sido considerablemente amplificada y es muy activa durante las primeras horas de la fermentación. Cuando se utiliza ácido ricinoleico como sustrato de fermentación en lugar de HOSFFA, se detectan altos niveles de ácidos grasos  $\omega$ -hidroxi (con relación a la cantidad de diácido producida) durante la fermentación. Por esta razón, las fermentaciones con ácido ricinoleico son una prueba más definitiva de la efectividad de la amplificación de *FAO1*. Cuando se compararon las fermentaciones de HDC40-1, HDC40-5 y HDC40-7, se descubrió que el número de copias del gen *FAO1* estaba correlacionado en forma inversa con el reducido nivel de ácidos grasos  $\omega$ -hidroxi, es decir el porcentaje de ácidos grasos  $\omega$ -hidroxi producidos con relación al nivel total de producto oxidado (Figura 10).

## 50 EJEMPLO 10

Identificación de motivos firma únicos para *FAO1* y *FAO2*.

Las secuencias de aminoácidos correspondientes a los genes objetivo *FAO1* y *FAO2* fueron examinados por la presencia de uno o más de las siete secuencias de péptidos identificadas por Slabas et al. (12) como indicativas de un gen *FAO*. La Tabla 7 muestra una comparación de los péptidos firma identificados por Slabas et al. entre los genes *FOA1* y *FAO2* de la presente invención, y el gen *FAOT* de *C. tropicalis*. También se comparan secuencias similares de la enzima *FAO* de *Candida albicans* (FAOCA). Como se refleja en la Tabla 7, todos los siete péptidos

FAOT están de acuerdo con los péptidos firma identificados por Slabas et al. Seis de los siete péptidos FAO1 pero únicamente cuatro de los siete péptidos FAO2 están de acuerdo con los péptidos firma. En forma muy interesante, aunque FAO2 es más cercano a FAOT en la identidad y similitud de la secuencia de aminoácidos (Tabla 3), FAO1 es más similar a FAOT cuando se comparan los siete péptidos firma.

- 5 El péptido 4 de FAO1 y de FAO2 (CGFCYLGC) es un péptido único, no encontrado en las FAO previamente caracterizadas. El péptido 1 de FAO2a y FAO2b (IIGSGAGAGVMA) así como el péptido 4 tanto de FAO1 [como de] FAO2a y FAO2b (CGFCYLGC) es también un péptido único, no encontrado en las FAO previamente caracterizadas.

Tabla 7. Comparación de los péptidos firma de FAO1 y FAO2 de Cognis con FAOT de *C. tropicalis* y FAOCA de *C. albicans*

	Péptido 1	Péptido 2	Péptido 3	Péptido 4	Péptido 5	Péptido 6	Péptido 7
Péptido firma <sup>1,2</sup>	IIGSG(X)GAGVVA (SEQ ID NO: 15)	AGSTFGGG (SEQ ID NO: 18)	NWSACLKTP (SEQ ID NO: 20)	CG(X)CHLGC (SEQ ID NO: 22)	IG(X)NL(X)LHPVS (SEQ ID NO: 24)	SAHQMS(X)CRMSG (SEQ ID NO: 27)	PTASG(X)NPM (SEQ ID NO: 30)
FAOT	IIGSGAGAGVVA (SEQ ID NO: 16)	AGSTFGGG (SEQ ID NO: 18)	NWSACLKTP (SEQ ID NO: 20)	CGFCHLGC (SEQ ID NO: 23)	IGKNLTLHPVS (SEQ ID NO: 25)	SAHQMSTCRMSG (SEQ ID NO: 28)	PTASGANPM (SEQ ID NO: 31)
FAO1	IIGSGAGAGVVA (SEQ ID NO: 16)	AGSTFGGG (SEQ ID NO: 18)	NWSACLKTP (SEQ ID NO: 20)	CGFCYLGC (SEQ ID NO: 13)	IGKNLTLHPVS (SEQ ID NO: 25)	SAHQMSSCRMSG (SEQ ID NO: 29)	PTASGANPM (SEQ ID NO: 31)
FAO2a	IIGSGAGAGVMA (SEQ ID NO: 14)	AGSTLGGG (SEQ ID NO: 19)	NWSACLKTP (SEQ ID NO: 20)	CGFCYLGC (SEQ ID NO: 13)	IGKNLTLHPVS (SEQ ID NO: 25)	SAHQMSTCRMSG (SEQ ID NO: 28)	PTASGANPM (SEQ ID NO: 31)
FAO2b	IIGSGAGAGVMA (SEQ ID NO: 14)	AGSTLGGG (SEQ ID NO: 19)	NWSACLKTP (SEQ ID NO: 20)	CGFCYLGC (SEQ ID NO: 13)	IGKNLTLHPVS (SEQ ID NO: 25)	SAHQMSTCRMSG (SEQ ID NO: 28)	PTASGANPM (SEQ ID NO: 31)
FAOCA	IIGSGAGSGVVA (SEQ ID NO: 17)	AGSTFGGG (SEQ ID NO: 18)	NWSACIKTP. (SEQ ID NO: 21)	CGFCHLGC (SEQ ID NO: 23)	IGANLTLHPVT (SEQ ID NO: 26)	SAHQMSSCRMSG (SEQ ID NO: 29)	PTASGANPM (SEQ ID NO: 31)

1 Los aminoácidos en negrilla indican las diferencias de aminoácidos con respecto al péptido firma

2(X) indica que cualquier aminoácido puede estar en esta posición



LISTADO DE SECUENCIAS

<110> Eirich, L. Dudley Craft, David L.

<120> Genes y proteínas de oxidasa de alcohol graso de *Candida tropicalis* y métodos relacionados con los mismos

<130> U0154 (1010-92)

<150> 60/374.021

<151> 2002-04-19

<160> 72

<170> PatentIn versión 3.2

<210> 1

<211> 4296

<212> ADN

<213> *Candida tropicalis*

<220>

<221> CDS

<222> (1940)..(4051)

<400> 1

ES 2 379 776 T3

tgcaactcgc	gagcatatcg	ccatcgtcca	tatcgttggc	actccatcca	ctgagtcagc	60
caagaagcaa	ttgttggtgc	accacacctt	aggtaatggt	gactttactg	ttttccacaa	120
gatctcgtca	ttcatcagtg	ccactactgc	tgggttgacc	gaccagaca	ccgccgctga	180
tgaaattgat	agagtgattg	agtcagccta	catcaaccag	cgtccaacgt	acttgggatt	240
cccttccaac	atggttgacg	ttcaagtgcc	agtcagcaag	ttggacaagc	cattgaactt	300
aacccacct	gcaaacaatc	caaagatcca	gtctgaggtc	ttgagcgaca	ttattgcctt	360
gattgaaacc	gccaaggatc	cagttatcat	cattgatgct	tgttgaggaa	ggcacaatgc	420
tacccagag	gcacagaagt	tgattgagtt	gacaaagttc	aagtttgctg	tcacccaat	480
ggctaaagg	tctaaggaca	ttgatgaaag	tgatccgaag	ttcattgggt	gctacgttgg	540
tgacttgtct	tatccaagag	tcaaagagtt	ggttgaaagc	tcggacttgg	tcttgtcctt	600
gggtgctgct	ttgtctgatt	tcaacactgg	ttcgttctca	tactcttgg	acaatgcca	660
ggttggtgaa	ttcactccg	actacactca	aatcaagagc	gctcagtacc	caggtatcag	720
aatgaaggaa	ttgttgggca	agttggttga	ggagccagaa	ttggtcaaga	cgtggttcaa	780
gatcccagca	aagaagttgg	tactgacaa	ctttgaacca	ttggtcttgc	caccggacca	840
caagctcacc	caatcctggt	tgtggagtaa	cttgggtaat	tggttgaaag	aaggtgatgt	900

gattgttacc gaaaccggta cttccaatth cggatttgte cagaccaaht tcccaaagaa 960  
tgctgtcggd atctcgcaag tcttgtgggg ttccattggc tactcggtecg gttctgccgc 1020  
tggtgccgth atcgccgccc aggagcttga tcccagccgt agagtcattct tgtttgttgg 1080  
tgacggttct ttgcagttga ccgtgcagga aatctccacc atggccagac acaagaacaa 1140  
catctacatc tttgtcttga acaacaacgg tttcaccatt gaaagattga ttcacggtecc 1200  
agaagctggt tacaacagta ttcaagaatg ggagaacgct gagttattga agactttcaa 1260  
ggctaccaac tacgagagtt tcaccgtcaa gactgtcggc gaacttgaca aggtgttcaa 1320  
ggatgaaaag tttgccgtca acgacaagat tagattggth gagatcatgt tagacactth 1380  
cgatgctcca gagaacttgg ttaagcaagc tgagagatct gccaacacca acaagtagag 1440  
tttgtctatg ttttccgtht gcctthtctt tctagtacga gacgttattg aacgaagtht 1500  
ttatatatct agatctaata catattccat gtctgttcat ttttgacgga gtttcataag 1560  
gtggcagtht ctaatcaaag gtccgtcatt ggcgtcgtgg cattggccggc tcccatcaac 1620  
tcgtatgtca atattttctg ttaactccgc cagacatacg atcaaaacct acaagcaaaa 1680  
aaattccaca tgctttgtth gagatctcca caacaacaa cggggtaaga aaatcatggg 1740  
gcgattaatc atgccatctt tgtaaatttc tttgtttcaa catcaccctc tttagtcaaa 1800  
ccttcacagg actgtctgct ctactttgcc acccagttca tatataaatt accaacttcc 1860  
accgagcacc accaacacct caccacctc tctcccccc cttttttttt ccagcttaga 1920  
cacacacttc aaactcgac atg gct cca ttt ttg ccc gac cag gtc gac tac 1972  
Met Ala Pro Phe Leu Pro Asp Gln Val Asp Tyr  
1 5 10  
aaa cac gtc gac acc ctt atg tta tta tgt gac ggg atc atc cac gaa 2020  
Lys His Val Asp Thr Leu Met Leu Leu Cys Asp Gly Ile Ile His Glu  
15 20 25  
acc acc gtg gac gaa atc aaa gac gtc att gcc cct gac ttc ccc gcc 2068  
Thr Thr Val Asp Glu Ile Lys Asp Val Ile Ala Pro Asp Phe Pro Ala  
30 35 40  
gac aaa tac gag gag tac gtc agg aca ttc acc aaa ccc tcc gaa acc 2116  
Asp Lys Tyr Glu Glu Tyr Val Arg Thr Phe Thr Lys Pro Ser Glu Thr  
45 50 55  
cca ggg ttc agg gaa acc gtc tac aac acc gtc aac gca aac acc atg 2164  
Pro Gly Phe Arg Glu Thr Val Tyr Asn Thr Val Asn Ala Asn Thr Met  
60 65 70 75  
gat gca atc cac cag ttc att atc ttg acc aat gtt ttg gga tca agg 2212  
Asp Ala Ile His Gln Phe Ile Ile Leu Thr Asn Val Leu Gly Ser Arg

ES 2 379 776 T3

				80					85					90		
gtc	ttg	gca	cca	gct	ttg	acc	aac	tcg	ttg	act	cct	atc	aag	gac	atg	2260
Val	Leu	Ala	Pro	Ala	Leu	Thr	Asn	Ser	Leu	Thr	Pro	Ile	Lys	Asp	Met	
			95					100					105			
agc	ttg	gaa	gac	cgt	gaa	aag	ttg	tta	gcc	tcg	tgg	cgt	gac	tcc	cct	2308
Ser	Leu	Glu	Asp	Arg	Glu	Lys	Leu	Leu	Ala	Ser	Trp	Arg	Asp	Ser	Pro	
		110					115					120				
att	gct	gct	aaa	agg	aag	ttg	ttc	agg	ttg	ggt	tct	acg	ctt	acc	ttg	2356
Ile	Ala	Ala	Lys	Arg	Lys	Leu	Phe	Arg	Leu	Val	Ser	Thr	Leu	Thr	Leu	
		125				130					135					
gtc	acg	ttc	acg	aga	ttg	gcc	aat	gag	ttg	cat	ttg	aaa	gcc	att	cat	2404
Val	Thr	Phe	Thr	Arg	Leu	Ala	Asn	Glu	Leu	His	Leu	Lys	Ala	Ile	His	
					145					150					155	
tat	cca	gga	aga	gaa	gac	cgt	gaa	aag	gct	tat	gaa	acc	cag	gag	att	2452
Tyr	Pro	Gly	Arg	Glu	Asp	Arg	Glu	Lys	Ala	Tyr	Glu	Thr	Gln	Glu	Ile	
				160					165					170		
gac	cct	ttt	aag	tac	cag	ttt	ttg	gaa	aaa	ccg	aag	ttt	tac	ggc	gct	2500
Asp	Pro	Phe	Lys	Tyr	Gln	Phe	Leu	Glu	Lys	Pro	Lys	Phe	Tyr	Gly	Ala	
			175					180					185			
gag	ttg	tac	ttg	cca	gat	att	gat	gtg	atc	att	att	gga	tct	ggg	gcc	2548
Glu	Leu	Tyr	Leu	Pro	Asp	Ile	Asp	Val	Ile	Ile	Ile	Gly	Ser	Gly	Ala	
		190					195					200				
ggt	gct	ggt	gtc	gtg	gcc	cac	act	ttg	acc	aac	gac	ggc	ttc	aag	agt	2596
Gly	Ala	Gly	Val	Val	Ala	His	Thr	Leu	Thr	Asn	Asp	Gly	Phe	Lys	Ser	
		205				210					215					
ttg	ggt	ttg	gaa	aag	ggc	aga	tac	ttt	agc	aac	tcc	gag	ttg	aac	ttt	2644
Leu	Val	Leu	Glu	Lys	Gly	Arg	Tyr	Phe	Ser	Asn	Ser	Glu	Leu	Asn	Phe	
		220			225					230					235	
gat	gac	aag	gac	ggg	ggt	caa	gaa	tta	tac	caa	agt	gga	ggt	act	ttg	2692
Asp	Asp	Lys	Asp	Gly	Val	Gln	Glu	Leu	Tyr	Gln	Ser	Gly	Gly	Thr	Leu	
				240				245						250		
acc	acc	gtc	aac	cag	cag	ttg	ttt	ggt	ctt	gct	ggt	tcc	act	ttt	ggt	2740
Thr	Thr	Val	Asn	Gln	Gln	Leu	Phe	Val	Leu	Ala	Gly	Ser	Thr	Phe	Gly	
			255					260					265			
ggt	ggt	acc	act	gtc	aat	tgg	tcg	gcc	tgt	ctt	aaa	acg	cca	ttc	aag	2788
Gly	Gly	Thr	Thr	Val	Asn	Trp	Ser	Ala	Cys	Leu	Lys	Thr	Pro	Phe	Lys	
		270					275					280				
gtg	cgt	aag	gaa	tgg	tat	gat	gag	ttt	ggc	ggt	gac	ttt	gct	gcc	gat	2836
Val	Arg	Lys	Glu	Trp	Tyr	Asp	Glu	Phe	Gly	Val	Asp	Phe	Ala	Ala	Asp	
		285				290					295					
gaa	gcc	tac	gac	aaa	gca	cag	gat	tat	ggt	tgg	cag	caa	atg	gga	gct	2884
Glu	Ala	Tyr	Asp	Lys	Ala	Gln	Asp	Tyr	Val	Trp	Gln	Gln	Met	Gly	Ala	
		300			305					310					315	

ES 2 379 776 T3

tct acc gaa ggc atc acc cac tct ttg gct aac gag att att att gaa	2932
Ser Thr Glu Gly Ile Thr His Ser Leu Ala Asn Glu Ile Ile Ile Glu	
320 325 330	
ggt ggc aag aaa tta ggt tac aag gcc aag gta tta gac caa aac agc	2980
Gly Gly Lys Lys Leu Gly Tyr Lys Ala Lys Val Leu Asp Gln Asn Ser	
335 340 345	
ggt ggt cat cct cat cac aga tgc ggt ttc tgt tat ttg ggt tgt aag	3028
Gly Gly His Pro His His Arg Cys Gly Phe Cys Tyr Leu Gly Cys Lys	
350 355 360	
cac ggt atc aag cag ggc tct gtt aat aac tgg ttt aga gac gca gct	3076
His Gly Ile Lys Gln Gly Ser Val Asn Asn Trp Phe Arg Asp Ala Ala	
365 370 375	
gcc cac ggt tct cag ttc atg caa cag gtt aga gtt ttg caa atc ctt	3124
Ala His Gly Ser Gln Phe Met Gln Gln Val Arg Val Leu Gln Ile Leu	
380 385 390 395	
aac aag aag ggc atc gct tat ggt atc ttg tgt gag gat gtt gta acc	3172
Asn Lys Lys Gly Ile Ala Tyr Gly Ile Leu Cys Glu Asp Val Val Thr	
400 405 410	
ggt gcc aag ttc acc att act ggc ccc aaa aag ttt gtt gtt gcc gcc	3220
Gly Ala Lys Phe Thr Ile Thr Gly Pro Lys Lys Phe Val Val Ala Ala	
415 420 425	
ggc gcc tta aac act cca tct gtg ttg gtc aac tcc gga ttc aag aac	3268
Gly Ala Leu Asn Thr Pro Ser Val Leu Val Asn Ser Gly Phe Lys Asn	
430 435 440	
aag aac atc ggt aag aac tta act ttg cat cca gtt tct gtc gtg ttt	3316
Lys Asn Ile Gly Lys Asn Leu Thr Leu His Pro Val Ser Val Val Phe	
445 450 455	
ggt gat ttt ggc aaa gac gtt caa gca gat cac ttc cac aac tcc atc	3364
Gly Asp Phe Gly Lys Asp Val Gln Ala Asp His Phe His Asn Ser Ile	
460 465 470 475	
atg act gct ctt tgt tca gaa gcc gct gat tta gac ggc aag ggt cat	3412
Met Thr Ala Leu Cys Ser Glu Ala Ala Asp Leu Asp Gly Lys Gly His	
480 485 490	
gga tgc aga att gaa acc atc ttg aac gct cca ttc atc cag gct tca	3460
Gly Cys Arg Ile Glu Thr Ile Leu Asn Ala Pro Phe Ile Gln Ala Ser	
495 500 505	
ttc tta cca tgg aga ggt agt aac gag gct aga cga gac ttg ttg cgt	3508
Phe Leu Pro Trp Arg Gly Ser Asn Glu Ala Arg Arg Asp Leu Leu Arg	
510 515 520	
tac aac aac atg gtg gcc atg tta ctt ctt agt cgt gat acc acc agt	3556
Tyr Asn Asn Met Val Ala Met Leu Leu Leu Ser Arg Asp Thr Thr Ser	
525 530 535	

ES 2 379 776 T3

ggt tcc gtt tcg tcc cat cca act aaa cct gaa gca tta gtt gtc gag	3604
Gly Ser Val Ser Ser His Pro Thr Lys Pro Glu Ala Leu Val Val Glu	
540 545 550 555	
tac gac gtg aac aag ttt gac aga aac tcc atc ttg cag gca ttg ttg	3652
Tyr Asp Val Asn Lys Phe Asp Arg Asn Ser Ile Leu Gln Ala Leu Leu	
560 565 570	
gtc act gct gac ttg ttg tac att caa ggt gcc aag aga atc ctt agt	3700
Val Thr Ala Asp Leu Leu Tyr Ile Gln Gly Ala Lys Arg Ile Leu Ser	
575 580 585	
ccc caa cca tgg gtg cca att ttt gaa tcc gac aag cca aag gat aag	3748
Pro Gln Pro Trp Val Pro Ile Phe Glu Ser Asp Lys Pro Lys Asp Lys	
590 595 600	
aga tca atc aag gac gag gac tat gtc gaa tgg aga gcc aag gtt gcc	3796
Arg Ser Ile Lys Asp Glu Asp Tyr Val Glu Trp Arg Ala Lys Val Ala	
605 610 615	
aag att cct ttt gac acc tac ggc tcg cct tat ggt tcg gcg cat caa	3844
Lys Ile Pro Phe Asp Thr Tyr Gly Ser Pro Tyr Gly Ser Ala His Gln	
620 625 630 635	
atg tct tct tgt cgt atg tca ggt aag ggt cct aaa tac ggt gct gtt	3892
Met Ser Ser Cys Arg Met Ser Gly Lys Gly Pro Lys Tyr Gly Ala Val	
640 645 650	
gat acc gat ggt aga ttg ttt gaa tgt tcg aat gtt tat gtt gct gac	3940
Asp Thr Asp Gly Arg Leu Phe Glu Cys Ser Asn Val Tyr Val Ala Asp	
655 660 665	
gct agt ctt ttg cca act gct agc ggt gct aat cct atg gtc acc acc	3988
Ala Ser Leu Leu Pro Thr Ala Ser Gly Ala Asn Pro Met Val Thr Thr	
670 675 680	
atg act ctt gca aga cat gtt gcg tta ggt ttg gca gac tcc ttg aag	4036
Met Thr Leu Ala Arg His Val Ala Leu Gly Leu Ala Asp Ser Leu Lys	
685 690 695	
acc aag gcc aag ttg tagttctgta tacgtatcctt ataatttaga tttcctttta	4091
Thr Lys Ala Lys Leu	
700	
ttgacggtaa acattcagga taggtactac ccttgctgca aaagcccagc acgcccgaat	4151
cgcgatgact tgagcgaagc aaacacgcac acaaaagggg tacacaaaaa ataacgagat	4211
gcccttgaag cacacaccca aacacgatgg aacacaagat ggccttagaa agtacaaaaa	4271
aagtaaagcc acttgattcc gccca	4296

<210> 2

<211> 704

<212> PRT

<213> *Candida tropicalis*

<400> 2

ES 2 379 776 T3

Met Ala Pro Phe Leu Pro Asp Gln Val Asp Tyr Lys His Val Asp Thr  
1 5 10 15

Leu Met Leu Leu Cys Asp Gly Ile Ile His Glu Thr Thr Val Asp Glu  
20 25 30

Ile Lys Asp Val Ile Ala Pro Asp Phe Pro Ala Asp Lys Tyr Glu Glu  
35 40 45

Tyr Val Arg Thr Phe Thr Lys Pro Ser Glu Thr Pro Gly Phe Arg Glu  
50 55 60

Thr Val Tyr Asn Thr Val Asn Ala Asn Thr Met Asp Ala Ile His Gln  
65 70 75 80

Phe Ile Ile Leu Thr Asn Val Leu Gly Ser Arg Val Leu Ala Pro Ala  
85 90 95

Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met Ser Leu Glu Asp Arg  
100 105 110

Glu Lys Leu Leu Ala Ser Trp Arg Asp Ser Pro Ile Ala Ala Lys Arg  
115 120 125

Lys Leu Phe Arg Leu Val Ser Thr Leu Thr Leu Val Thr Phe Thr Arg  
130 135 140

Leu Ala Asn Glu Leu His Leu Lys Ala Ile His Tyr Pro Gly Arg Glu  
145 150 155 160

Asp Arg Glu Lys Ala Tyr Glu Thr Gln Glu Ile Asp Pro Phe Lys Tyr  
165 170 175

Gln Phe Leu Glu Lys Pro Lys Phe Tyr Gly Ala Glu Leu Tyr Leu Pro  
180 185 190

Asp Ile Asp Val Ile Ile Ile Gly Ser Gly Ala Gly Ala Gly Val Val  
195 200 205

Ala His Thr Leu Thr Asn Asp Gly Phe Lys Ser Leu Val Leu Glu Lys  
210 215 220



ES 2 379 776 T3

Gly Arg Tyr Phe Ser Asn Ser Glu Leu Asn Phe Asp Asp Lys Asp Gly  
 225 230 235 240

Val Gln Glu Leu Tyr Gln Ser Gly Gly Thr Leu Thr Thr Val Asn Gln  
 245 250 255

Gln Leu Phe Val Leu Ala Gly Ser Thr Phe Gly Gly Gly Thr Thr Val  
 260 265 270

Asn Trp Ser Ala Cys Leu Lys Thr Pro Phe Lys Val Arg Lys Glu Trp  
 275 280 285

Tyr Asp Glu Phe Gly Val Asp Phe Ala Ala Asp Glu Ala Tyr Asp Lys  
 290 295 300

Ala Gln Asp Tyr Val Trp Gln Gln Met Gly Ala Ser Thr Glu Gly Ile  
 305 310 315 320

Thr His Ser Leu Ala Asn Glu Ile Ile Ile Glu Gly Gly Lys Lys Leu  
 325 330 335

Gly Tyr Lys Ala Lys Val Leu Asp Gln Asn Ser Gly Gly His Pro His  
 340 345 350

His Arg Cys Gly Phe Cys Tyr Leu Gly Cys Lys His Gly Ile Lys Gln  
 355 360 365

Gly Ser Val Asn Asn Trp Phe Arg Asp Ala Ala Ala His Gly Ser Gln  
 370 375 380

Phe Met Gln Gln Val Arg Val Leu Gln Ile Leu Asn Lys Lys Gly Ile  
 385 390 395 400

Ala Tyr Gly Ile Leu Cys Glu Asp Val Val Thr Gly Ala Lys Phe Thr  
 405 410 415

Ile Thr Gly Pro Lys Lys Phe Val Val Ala Ala Gly Ala Leu Asn Thr  
 420 425 430

Pro Ser Val Leu Val Asn Ser Gly Phe Lys Asn Lys Asn Ile Gly Lys  
 435 440 445

ES 2 379 776 T3

Asn Leu Thr Leu His Pro Val Ser Val Val Phe Gly Asp Phe Gly Lys  
 450 455 460

Asp Val Gln Ala Asp His Phe His Asn Ser Ile Met Thr Ala Leu Cys  
 465 470 475 480

Ser Glu Ala Ala Asp Leu Asp Gly Lys Gly His Gly Cys Arg Ile Glu  
 485 490 495

Thr Ile Leu Asn Ala Pro Phe Ile Gln Ala Ser Phe Leu Pro Trp Arg  
 500 505 510

Gly Ser Asn Glu Ala Arg Arg Asp Leu Leu Arg Tyr Asn Asn Met Val  
 515 520 525

Ala Met Leu Leu Leu Ser Arg Asp Thr Thr Ser Gly Ser Val Ser Ser  
 530 535 540

His Pro Thr Lys Pro Glu Ala Leu Val Val Glu Tyr Asp Val Asn Lys  
 545 550 555 560

Phe Asp Arg Asn Ser Ile Leu Gln Ala Leu Leu Val Thr Ala Asp Leu  
 565 570 575

Leu Tyr Ile Gln Gly Ala Lys Arg Ile Leu Ser Pro Gln Pro Trp Val  
 580 585 590

Pro Ile Phe Glu Ser Asp Lys Pro Lys Asp Lys Arg Ser Ile Lys Asp  
 595 600 605

Glu Asp Tyr Val Glu Trp Arg Ala Lys Val Ala Lys Ile Pro Phe Asp  
 610 615 620

Thr Tyr Gly Ser Pro Tyr Gly Ser Ala His Gln Met Ser Ser Cys Arg  
 625 630 635 640

Met Ser Gly Lys Gly Pro Lys Tyr Gly Ala Val Asp Thr Asp Gly Arg  
 645 650 655

Leu Phe Glu Cys Ser Asn Val Tyr Val Ala Asp Ala Ser Leu Leu Pro  
 660 665 670

ES 2 379 776 T3

Thr Ala Ser Gly Ala Asn Pro Met Val Thr Thr Met Thr Leu Ala Arg  
675 680 685

His Val Ala Leu Gly Leu Ala Asp Ser Leu Lys Thr Lys Ala Lys Leu  
690 695 700

<210> 3

<211> 4158

<212> ADN

<213> *Candida tropicalis*

<220>

<221> CDS

<222> (1521)..(3632)

<400> 3

ES 2 379 776 T3

tctaccaag tacgagaacg agcttggtga tagagcttag acttgtcttt tgtatttgta	60
atctgacggt gaccgtttga gtttttctctg tgatatcacg taaatctggc aaccagcttt	120
ctatTTTTTT tgcaacactt tttctcttca ccactctcag aaccaatgcc accgaagaag	180
ggctcaacc aggaggaaaa gctctcgtca atcctcgcct ggttccaaag cagccactgc	240
ttctatacgc tcaaagaggt tgaacagaag gctagcaaag cgtgcaagat ctcgtctatg	300
cagatcaagg acttggttgc aaccttagtc aacgaagggt tagtggaaca ggaaaaatgt	360
gggaccacca acttgtactg gtcgttcccg tactcggaac acaaacgcaa gctacagaga	420
tacgagcagc taagacaatc cgttgccaaa cttcaagcga ataaaggcaa gttggcggaa	480
gagttgcgaa acgcgtgtgg tgagcgtgac atggacagca ataggctaaa ccggatgcaa	540
cagtgcgac agcttgttca cgaggcggca cgctccagg aggaactaaa actgtcgagg	600
cagagagata ccattgacga gttggttcag gccattgact tcttcaacga gctgatagag	660
accgtcctca gctacatcag ccatcagtcg gggaccagcg tgtcggtatt gaaaacggag	720
ttgagatac ccgcagaact agaagaggct ccccagataa acaatgccgg agttagtgcg	780
taaatcgagc atgcatacgt tggagagaaa tagagaaaca gatttccggt gaaacgctac	840
aacacagacg aggaatacag aatggaacat gacggaaata taatatccga ggaaagacga	900
aagtacgaca tggaactccg ttactgcaac atcgatcgtg ctagatacga catagaacia	960
tgttgctatt acatggaaag ctggtgctac aatccagaat acggttgtag tcaagggaga	1020
tgaggctggg agccgagtgg tacataaata ggcatatagg accgtcactt ggtctaggat	1080
cgtgtagagg gtggaagagg taggcaagat ccattctaact ctactgagtg acggctaata	1140

ES 2 379 776 T3

tacgatcagc gttctcaggc gagcacagtc attcctcatt tctgtacata cgttgcccct	1200
ttatgttttt tttcacagga tgctcacgcc caacatttcc cccacattt tattaccac	1260
attgagccgt caaatgcatt ttttttatcc gtcgcttgct aagacaaaat tccacatgct	1320
ttgtctcaga gtatataaac aacggggcaa aaaaacatgg ggtaaatagc ttattcgtgg	1380
attgatattt ttatatttta gttcgcccct ttcgccacca agctcaattg gactatttgt	1440
cagtggtgta taagctagag attactagac tgcttttctg attccttgatt ttccttttc	1500
attagttcca gtacctagag atg aat acc ttc ttg cca gac gtg ctc gaa tac	1553
Met Asn Thr Phe Leu Pro Asp Val Leu Glu Tyr	
1 5 10	
aaa cac gtc gac acc ctt ttg tta ttg tgt gac ggg atc atc cac gaa	1601
Lys His Val Asp Thr Leu Leu Leu Leu Cys Asp Gly Ile Ile His Glu	
15 20 25	
acc aca gtc gat cag atc aag gac gcc att gct ccc gac ttc cct gag	1649
Thr Thr Val Asp Gln Ile Lys Asp Ala Ile Ala Pro Asp Phe Pro Glu	
30 35 40	
gac cag tac gag gag tat ctc aag acc ttc acc aag cca tct gag acc	1697
Asp Gln Tyr Glu Glu Tyr Leu Lys Thr Phe Thr Lys Pro Ser Glu Thr	
45 50 55	
cct ggg ttc aga gaa gcc gtc tac gac acg atc aac gcc acc cca acc	1745
Pro Gly Phe Arg Glu Ala Val Tyr Asp Thr Ile Asn Ala Thr Pro Thr	
60 65 70 75	
gat gcc gtg cac atg tgt att gtc ttg acc acc gca ttg gac tcc aga	1793
Asp Ala Val His Met Cys Ile Val Leu Thr Thr Ala Leu Asp Ser Arg	
80 85 90	
atc ttg gcc ccc acg ttg acc aac tcg ttg acg cct atc aag gat atg	1841
Ile Leu Ala Pro Thr Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met	
95 100 105	
acc ttg aag gag cgt gaa caa ttg ttg gcc tct tgg cgt gat tcc ccg	1889
Thr Leu Lys Glu Arg Glu Gln Leu Leu Ala Ser Trp Arg Asp Ser Pro	
110 115 120	
att gcg gca aag aga aga ttg ttc aga ttg att tcc tcg ctt acc ttg	1937
Ile Ala Ala Lys Arg Arg Leu Phe Arg Leu Ile Ser Ser Leu Thr Leu	
125 130 135	
acg acg ttt acg aga ttg gcc agc gaa ttg cac ttg aaa gcc atc cac	1985
Thr Thr Phe Thr Arg Leu Ala Ser Glu Leu His Leu Lys Ala Ile His	
140 145 150 155	
tac cct ggc aga gac ttg cgt gaa aag gcg tat gaa acc cag gtg gtt	2033
Tyr Pro Gly Arg Asp Leu Arg Glu Lys Ala Tyr Glu Thr Gln Val Val	
160 165 170	
gac cct ttc agg tac ctg ttt atg gag aaa cca aag ttt gac gcc gcc	2081

ES 2 379 776 T3

Asp	Pro	Phe	Arg	Tyr	Leu	Phe	Met	Glu	Lys	Pro	Lys	Phe	Asp	Gly	Ala		
			175					180					185				
gaa	ttg	tac	ttg	cca	gat	atc	gac	gtc	atc	atc	att	gga	tca	ggc	gcc		2129
Glu	Leu	Tyr	Leu	Pro	Asp	Ile	Asp	Val	Ile	Ile	Ile	Gly	Ser	Gly	Ala		
		190					195					200					
ggt	gct	ggt	gtc	atg	gcc	cac	act	ctc	gcc	aac	gac	ggg	ttc	aag	acc		2177
Gly	Ala	Gly	Val	Met	Ala	His	Thr	Leu	Ala	Asn	Asp	Gly	Phe	Lys	Thr		
	205					210					215						
ttg	ggt	ttg	gaa	aag	gga	aag	tat	ttc	agc	aac	tcc	gag	ttg	aac	ttt		2225
Leu	Val	Leu	Glu	Lys	Gly	Lys	Tyr	Phe	Ser	Asn	Ser	Glu	Leu	Asn	Phe		
220				225						230					235		
aat	gac	gct	gat	ggc	gtg	aaa	gag	ttg	tac	caa	ggt	aaa	ggt	gct	ttg		2273
Asn	Asp	Ala	Asp	Gly	Val	Lys	Glu	Leu	Tyr	Gln	Gly	Lys	Gly	Ala	Leu		
				240					245					250			
gcc	acc	acc	aat	cag	cag	atg	ttt	att	ctt	gcc	ggt	tcc	act	ttg	ggc		2321
Ala	Thr	Thr	Asn	Gln	Gln	Met	Phe	Ile	Leu	Ala	Gly	Ser	Thr	Leu	Gly		
			255					260					265				
ggt	ggt	acc	act	gtc	aac	tgg	tct	gct	tgc	ctt	aaa	aca	cca	ttt	aaa		2369
Gly	Gly	Thr	Thr	Val	Asn	Trp	Ser	Ala	Cys	Leu	Lys	Thr	Pro	Phe	Lys		
		270					275					280					
gtg	cgt	aag	gag	tgg	tac	gac	gag	ttt	ggt	ctt	gaa	ttt	gct	gcc	gat		2417
Val	Arg	Lys	Glu	Trp	Tyr	Asp	Glu	Phe	Gly	Leu	Glu	Phe	Ala	Ala	Asp		
	285					290					295						
gaa	gcc	tac	gac	aaa	gcg	cag	gat	tat	ggt	tgg	aaa	caa	atg	ggt	gct		2465
Glu	Ala	Tyr	Asp	Lys	Ala	Gln	Asp	Tyr	Val	Trp	Lys	Gln	Met	Gly	Ala		
300				305						310					315		
tca	aca	gat	gga	atc	act	cac	tcc	ttg	gcc	aac	gaa	ggt	gtg	ggt	gaa		2513
Ser	Thr	Asp	Gly	Ile	Thr	His	Ser	Leu	Ala	Asn	Glu	Val	Val	Val	Glu		
				320				325						330			
gga	ggt	aag	aag	ttg	ggc	tac	aag	agc	aag	gaa	att	gag	cag	aac	aac		2561
Gly	Gly	Lys	Lys	Leu	Gly	Tyr	Lys	Ser	Lys	Glu	Ile	Glu	Gln	Asn	Asn		
		335						340					345				
ggt	ggc	cac	cct	gac	cac	cca	tgt	ggt	ttc	tgt	tac	ttg	ggc	tgt	aag		2609
Gly	Gly	His	Pro	Asp	His	Pro	Cys	Gly	Phe	Cys	Tyr	Leu	Gly	Cys	Lys		
		350					355					360					
tac	ggt	att	aaa	cag	ggt	tct	gtg	aat	aac	tgg	ttt	aga	gac	gca	gct		2657
Tyr	Gly	Ile	Lys	Gln	Gly	Ser	Val	Asn	Asn	Trp	Phe	Arg	Asp	Ala	Ala		
	365					370					375						
gcc	cac	ggg	tcc	aag	ttc	atg	caa	caa	gtc	aga	ggt	gtg	caa	atc	ctc		2705
Ala	His	Gly	Ser	Lys	Phe	Met	Gln	Gln	Val	Arg	Val	Val	Gln	Ile	Leu		
380				385					390						395		
aac	aag	aat	ggc	gtc	gct	tat	ggt	atc	ttg	tgt	gag	gat	gtc	gaa	acc		2753
Asn	Lys	Asn	Gly	Val	Ala	Tyr	Gly	Ile	Leu	Cys	Glu	Asp	Val	Glu	Thr		

ES 2 379 776 T3

				400					405					410					
gga	gtc	agg	ttc	act	att	agt	ggc	ccc	aaa	aag	ttt	gtt	gtt	tct	gct		2801		
Gly	Val	Arg	Phe	Thr	Ile	Ser	Gly	Pro	Lys	Lys	Phe	Val	Val	Ser	Ala				
				415					420					425					
ggt	tct	ttg	aac	acg	cca	act	gtg	ttg	acc	aac	tcc	gga	ttc	aag	aac		2849		
Gly	Ser	Leu	Asn	Thr	Pro	Thr	Val	Leu	Thr	Asn	Ser	Gly	Phe	Lys	Asn				
				430					435					440					
aag	cac	att	ggt	aag	aac	ttg	acg	ttg	cac	cca	ggt	tcc	acc	gtg	ttt		2897		
Lys	His	Ile	Gly	Lys	Asn	Leu	Thr	Leu	His	Pro	Val	Ser	Thr	Val	Phe				
				445					450					455					
ggt	gac	ttt	ggc	aga	gac	gtg	caa	gcc	gac	cat	ttc	cac	aaa	tct	att		2945		
Gly	Asp	Phe	Gly	Arg	Asp	Val	Gln	Ala	Asp	His	Phe	His	Lys	Ser	Ile				
				460					465					470					475
atg	act	tcg	ctt	tgt	tac	gag	ggt	gct	gac	ttg	gac	ggc	aag	ggc	cac		2993		
Met	Thr	Ser	Leu	Cys	Tyr	Glu	Val	Ala	Asp	Leu	Asp	Gly	Lys	Gly	His				
				480					485					490					
gga	tgc	aga	atc	gaa	acc	atc	ttg	aac	gct	cca	ttc	atc	caa	gct	tct		3041		
Gly	Cys	Arg	Ile	Glu	Thr	Ile	Leu	Asn	Ala	Pro	Phe	Ile	Gln	Ala	Ser				
				495					500					505					
ttg	ttg	cca	tgg	aga	gga	agt	gac	gag	gtc	aga	aga	gac	ttg	ttg	cgt		3089		
Leu	Leu	Pro	Trp	Arg	Gly	Ser	Asp	Glu	Val	Arg	Arg	Asp	Leu	Leu	Arg				
				510					515					520					
tac	aac	aac	atg	gtg	gcc	atg	ttg	ctt	atc	acg	cgt	gat	acc	acc	agt		3137		
Tyr	Asn	Asn	Met	Val	Ala	Met	Leu	Leu	Ile	Thr	Arg	Asp	Thr	Thr	Ser				
				525					530					535					
ggt	tca	ggt	tct	gct	gac	cca	aag	aag	ccc	gac	gct	ttg	att	gtc	gac		3185		
Gly	Ser	Val	Ser	Ala	Asp	Pro	Lys	Lys	Pro	Asp	Ala	Leu	Ile	Val	Asp				
				540					545					550					555
tat	gag	att	aac	aag	ttt	gac	aag	aat	gcc	atc	ttg	caa	gct	ttc	ttg		3233		
Tyr	Glu	Ile	Asn	Lys	Phe	Asp	Lys	Asn	Ala	Ile	Leu	Gln	Ala	Phe	Leu				
				560					565					570					
atc	act	tcc	gac	atg	ttg	tac	att	gaa	ggt	gcc	aag	aga	atc	ctc	agt		3281		
Ile	Thr	Ser	Asp	Met	Leu	Tyr	Ile	Glu	Gly	Ala	Lys	Arg	Ile	Leu	Ser				
				575					580					585					
cca	cag	cca	tgg	gtg	cca	atc	ttt	gag	tcg	aac	aag	cca	aag	gag	caa		3329		
Pro	Gln	Pro	Trp	Val	Pro	Ile	Phe	Glu	Ser	Asn	Lys	Pro	Lys	Glu	Gln				
				590					595					600					
aga	acg	atc	aag	gac	aag	gac	tat	ggt	gag	tgg	aga	gcc	aag	gct	gct		3377		
Arg	Thr	Ile	Lys	Asp	Lys	Asp	Tyr	Val	Glu	Trp	Arg	Ala	Lys	Ala	Ala				
				605					610					615					
aag	ata	cct	ttc	gac	acc	tac	ggt	tct	gca	tat	ggg	tcc	gca	cat	caa		3425		
Lys	Ile	Pro	Phe	Asp	Thr	Tyr	Gly	Ser	Ala	Tyr	Gly	Ser	Ala	His	Gln				
				620					625					630					635

ES 2 379 776 T3

atg tcc acc tgt cgt atg tcc gga aag ggt cct aaa tac ggt gct gtt	3473
Met Ser Thr Cys Arg Met Ser Gly Lys Gly Pro Lys Tyr Gly Ala Val	
640 645 650	
gat act gat ggt aga ttg ttt gaa tgt tcg aat gtc tat gtt gct gat	3521
Asp Thr Asp Gly Arg Leu Phe Glu Cys Ser Asn Val Tyr Val Ala Asp	
655 660 665	
gct agt gtt ttg cct act gcc agc ggt gcc aac cca atg ata tcc acc	3569
Ala Ser Val Leu Pro Thr Ala Ser Gly Ala Asn Pro Met Ile Ser Thr	
670 675 680	
atg acc ttt gct aga cag att gcg tta ggt ttg gct gac tcc ttg aag	3617
Met Thr Phe Ala Arg Gln Ile Ala Leu Gly Leu Ala Asp Ser Leu Lys	
685 690 695	
acc aaa ccc aag ttg tagagagacg gaaatacgac acttatatac tagatgtatc	3672
Thr Lys Pro Lys Leu	
700	
ttacaattta tattctcgat gatggctttt actatctcct atggttacct ataatgacat	3732
caccacaacc tctactactg tctccagtat cctccttgct gttgaccgta cccaccagcc	3792
tgttgattga accctgtgaa ctgtgggtgc tgttgagcgt accccacggt agtgaactgc	3852
ggttggtggg caaactgctg tacgggctgt tgctgctgct gctggtggtg ttggtggtgt	3912
tgccccgtgg gctggttgta caacgacatg atgttctgct tgtttgtctg ttgggcaacc	3972
aactgtgggt tattcatctg catcaactgc tgctgggtgt gagggttggt tggatccaag	4032
tactcttgcc cgttggcgtc gatataagaa atctgccccg tgactgggtc agtgtactgg	4092
tatatctgtg gcatgccacc agcttggtgca ggcatgccgg ttgccaatgg cacctgtgct	4152
tgcgtc	4158

<210> 4

<211> 704

<212> PRT

<213> *Candida tropicalis*

<400> 4



ES 2 379 776 T3

Met Asn Thr Phe Leu Pro Asp Val Leu Glu Tyr Lys His Val Asp Thr  
1 5 10 15

Leu Leu Leu Leu Cys Asp Gly Ile Ile His Glu Thr Thr Val Asp Gln  
20 25 30

Ile Lys Asp Ala Ile Ala Pro Asp Phe Pro Glu Asp Gln Tyr Glu Glu  
35 40 45

ES 2 379 776 T3

Tyr Leu Lys Thr Phe Thr Lys Pro Ser Glu Thr Pro Gly Phe Arg Glu  
 50 55 60

Ala Val Tyr Asp Thr Ile Asn Ala Thr Pro Thr Asp Ala Val His Met  
 65 70 75 80

Cys Ile Val Leu Thr Thr Ala Leu Asp Ser Arg Ile Leu Ala Pro Thr  
 85 90 95

Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met Thr Leu Lys Glu Arg  
 100 105 110

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

Arg Leu Phe Arg Leu Ile Ser Ser Leu Thr Leu Thr Thr Phe Thr Arg  
 130 135 140

Leu Ala Ser Glu Leu His Leu Lys Ala Ile His Tyr Pro Gly Arg Asp  
 145 150 155 160

Leu Arg Glu Lys Ala Tyr Glu Thr Gln Val Val Asp Pro Phe Arg Tyr  
 165 170 175

Leu Phe Met Glu Lys Pro Lys Phe Asp Gly Ala Glu Leu Tyr Leu Pro  
 180 185 190

Asp Ile Asp Val Ile Ile Ile Gly Ser Gly Ala Gly Ala Gly Val Met  
 195 200 205

Ala His Thr Leu Ala Asn Asp Gly Phe Lys Thr Leu Val Leu Glu Lys  
 210 215 220

Gly Lys Tyr Phe Ser Asn Ser Glu Leu Asn Phe Asn Asp Ala Asp Gly  
 225 230 235 240

Val Lys Glu Leu Tyr Gln Gly Lys Gly Ala Leu Ala Thr Thr Asn Gln  
 245 250 255

Gln Met Phe Ile Leu Ala Gly Ser Thr Leu Gly Gly Gly Thr Thr Val  
 260 265 270

ES 2 379 776 T3

Asn Trp Ser Ala Cys Leu Lys Thr Pro Phe Lys Val Arg Lys Glu Trp  
 275 280 285

Tyr Asp Glu Phe Gly Leu Glu Phe Ala Ala Asp Glu Ala Tyr Asp Lys  
 290 295 300

Ala Gln Asp Tyr Val Trp Lys Gln Met Gly Ala Ser Thr Asp Gly Ile  
 305 310 315 320

Thr His Ser Leu Ala Asn Glu Val Val Val Glu Gly Gly Lys Lys Leu  
 325 330 335

Gly Tyr Lys Ser Lys Glu Ile Glu Gln Asn Asn Gly Gly His Pro Asp  
 340 345 350

His Pro Cys Gly Phe Cys Tyr Leu Gly Cys Lys Tyr Gly Ile Lys Gln  
 355 360 365

Gly Ser Val Asn Asn Trp Phe Arg Asp Ala Ala Ala His Gly Ser Lys  
 370 375 380

Phe Met Gln Gln Val Arg Val Val Gln Ile Leu Asn Lys Asn Gly Val  
 385 390 395 400

Ala Tyr Gly Ile Leu Cys Glu Asp Val Glu Thr Gly Val Arg Phe Thr  
 405 410 415

Ile Ser Gly Pro Lys Lys Phe Val Val Ser Ala Gly Ser Leu Asn Thr  
 420 425 430

Pro Thr Val Leu Thr Asn Ser Gly Phe Lys Asn Lys His Ile Gly Lys  
 435 440 445

Asn Leu Thr Leu His Pro Val Ser Thr Val Phe Gly Asp Phe Gly Arg  
 450 455 460

Asp Val Gln Ala Asp His Phe His Lys Ser Ile Met Thr Ser Leu Cys  
 465 470 475 480

Tyr Glu Val Ala Asp Leu Asp Gly Lys Gly His Gly Cys Arg Ile Glu  
 485 490 495

ES 2 379 776 T3

Thr Ile Leu Asn Ala Pro Phe Ile Gln Ala Ser Leu Leu Pro Trp Arg  
 500 505 510

Gly Ser Asp Glu Val Arg Arg Asp Leu Leu Arg Tyr Asn Asn Met Val  
 515 520 525

Ala Met Leu Leu Ile Thr Arg Asp Thr Thr Ser Gly Ser Val Ser Ala  
 530 535 540

Asp Pro Lys Lys Pro Asp Ala Leu Ile Val Asp Tyr Glu Ile Asn Lys  
 545 550 555 560

Phe Asp Lys Asn Ala Ile Leu Gln Ala Phe Leu Ile Thr Ser Asp Met  
 565 570 575

Leu Tyr Ile Glu Gly Ala Lys Arg Ile Leu Ser Pro Gln Pro Trp Val  
 580 585 590

Pro Ile Phe Glu Ser Asn Lys Pro Lys Glu Gln Arg Thr Ile Lys Asp  
 595 600 605

Lys Asp Tyr Val Glu Trp Arg Ala Lys Ala Ala Lys Ile Pro Phe Asp  
 610 615 620

Thr Tyr Gly Ser Ala Tyr Gly Ser Ala His Gln Met Ser Thr Cys Arg  
 625 630 635 640

Met Ser Gly Lys Gly Pro Lys Tyr Gly Ala Val Asp Thr Asp Gly Arg  
 645 650 655

Leu Phe Glu Cys Ser Asn Val Tyr Val Ala Asp Ala Ser Val Leu Pro  
 660 665 670

Thr Ala Ser Gly Ala Asn Pro Met Ile Ser Thr Met Thr Phe Ala Arg  
 675 680 685

Gln Ile Ala Leu Gly Leu Ala Asp Ser Leu Lys Thr Lys Pro Lys Leu  
 690 695 700

<210> 5

<211> 3753

<212> ADN

<213> *Candida tropicalis*

<220>

<221> CDS

<222> (1099)..(3213)

<400> 5

ES 2 379 776 T3

tgttgataga gcttagactt gtgttttgta tttgtaatct gacgttgatc gtttgatatt	60
ttcctgtgat atcacgtaaa ttcggcaacc aactttttac tttttgcaac actttctctt	120
caccactctc agaaccaatg ccaccgaaga agggctctcag ccaggaggaa aagctctcgg	180
cactcctcac ctggttccaa gccagtcatt gcttctacac actcaaggag gttgaacaga	240
aggcgagcaa agcgtgcaag atctogtcta tgcagatcaa ggacttggtt gcaagcttag	300
tcaacgaagg tttggtagaa caggaaaagt gtgggaccac aaacttgtag tggtcgttcc	360
agtactcgga attcaaacgg aagctacaga gatacgggca gctaagacaa tcagccgcca	420
aacttcaagc ggataaaggc aagttggcgg aagagttgcg aaacgcatgt ggtgaacggg	480
acatggacaa caataggcaa gaccggatgc aacaatacga tcaccttggt aacgaggcgg	540
cacgtctcca ggaggaacta aaactgtcaa ggcagataga taccattgac gagttagttc	600
aggccattga tttcttcaac gagctgatag agaccgtcct cagctacatc agccatcagt	660
cagggaccag cgtgtcgata ttgaaaacgg agtttgagat acccgcagaa ctagaagagg	720
ccccccagat aagcaatgcc ggagtttagtg cgtaaatacga gcaggcatac attgcccctt	780
tgtatTTTTT cacaggatgc tcaccccacc acgcccaca tttccccca cattttatta	840
cccacattga gccgtcaaat gcattTTTTT atccgtcgtc agctaaacca aaattccaca	900
tgcgttgcct cagagtatat aaacaacggg gcaaaaaaca tgggattaat agcttatttg	960
tggattgata tttttatatt ttagttcgcc ccttctacga ccaagctcaa ttggactatt	1020
tgtcagtggt gtataagcta gagattacta gactgctttt ctgattcttg atcatccct	1080
tagttccagt gcctagag atg aat acc ttc ttg cca gac gtg ctc gaa tac	1131
Met Asn Thr Phe Leu Pro Asp Val Leu Glu Tyr	
1 5 10	
aaa cac gtc gat acc ctt ttg tta tta tgt gac ggg atc atc cac gaa	1179
Lys His Val Asp Thr Leu Leu Leu Leu Cys Asp Gly Ile Ile His Glu	
15 20 25	
acc aca gtc gac cag atc agg gac gcc att gct ccc gac ttc cct gaa	1227
Thr Thr Val Asp Gln Ile Arg Asp Ala Ile Ala Pro Asp Phe Pro Glu	
30 35 40	
gac cag tac gag gag tat ctc aag acc ttc acc aag cca tct gag acc	1275
Asp Gln Tyr Glu Glu Tyr Leu Lys Thr Phe Thr Lys Pro Ser Glu Thr	
45 50 55	

ES 2 379 776 T3

cct ggg ttc aga gaa gcc gtc tac gac acg atc aac agc acc cca acc	1323
Pro Gly Phe Arg Glu Ala Val Tyr Asp Thr Ile Asn Ser Thr Pro Thr	
60 65 70 75	
gag gct gtg cac atg tgt att gta ttg acc acc gca ttg gac tcg aga	1371
Glu Ala Val His Met Cys Ile Val Leu Thr Thr Ala Leu Asp Ser Arg	
80 85 90	
atc ttg gcc ccc acg ttg acc aac tcg ttg acg cct atc aag gat atg	1419
Ile Leu Ala Pro Thr Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met	
95 100 105	
acc ttg aaa gag cgt gaa caa ttg ttg gct gcc tgg cgt gat tcc ccg	1467
Thr Leu Lys Glu Arg Glu Gln Leu Leu Ala Ala Trp Arg Asp Ser Pro	
110 115 120	
atc gcg gcc aag aga aga ttg ttc aga ttg att tcc tca ctt acc ttg	1515
Ile Ala Ala Lys Arg Arg Leu Phe Arg Leu Ile Ser Ser Leu Thr Leu	
125 130 135	
acg acc ttt acg aga ttg gcc agc gac ttg cac ttg aga gcc atc cac	1563
Thr Thr Phe Thr Arg Leu Ala Ser Asp Leu His Leu Arg Ala Ile His	
140 145 150 155	
tac cct ggc aga gac ttg cgt gaa aag gca tat gaa acc cag gtg gtt	1611
Tyr Pro Gly Arg Asp Leu Arg Glu Lys Ala Tyr Glu Thr Gln Val Val	
160 165 170	
gac cct ttc agg tac ctg ttt atg gaa aaa cca aag ttt gac ggc acc	1659
Asp Pro Phe Arg Tyr Leu Phe Met Glu Lys Pro Lys Phe Asp Gly Thr	
175 180 185	
gag ttg tac ttg cca gat atc gac gtc atc atc att gga tcc ggt gcc	1707
Glu Leu Tyr Leu Pro Asp Ile Asp Val Ile Ile Ile Gly Ser Gly Ala	
190 195 200	
ggt gct ggt gtc atg gcc cac act tta gcc aac gac ggg tac aag acc	1755
Gly Ala Gly Val Met Ala His Thr Leu Ala Asn Asp Gly Tyr Lys Thr	
205 210 215	
ttg gtt ttg gaa aag gga aag tat ttc agc aac tcc gag ttg aac ttt	1803
Leu Val Leu Glu Lys Gly Lys Tyr Phe Ser Asn Ser Glu Leu Asn Phe	
220 225 230 235	
aat gat gcc gat ggt atg aaa gag ttg tac caa ggt aaa tgt gcg ttg	1851
Asn Asp Ala Asp Gly Met Lys Glu Leu Tyr Gln Gly Lys Cys Ala Leu	
240 245 250	
acc acc acg aac cag cag atg ttt att ctt gcc ggt tcc act ttg ggc	1899
Thr Thr Thr Asn Gln Gln Met Phe Ile Leu Ala Gly Ser Thr Leu Gly	
255 260 265	
ggt ggt acc act gtt aac tgg tct gct tgt ctt aaa aca cca ttt aaa	1947
Gly Gly Thr Thr Val Asn Trp Ser Ala Cys Leu Lys Thr Pro Phe Lys	
270 275 280	

ES 2 379 776 T3

gtg cgt aag gag tgg tac gac gag ttt ggt ctt gaa ttt gct gcc gac	1995
Val Arg Lys Glu Trp Tyr Asp Glu Phe Gly Leu Glu Phe Ala Ala Asp	
285 290 295	
gaa gcc tac gac aaa gca caa gac tat gtt tgg aaa caa atg ggc gct	2043
Glu Ala Tyr Asp Lys Ala Gln Asp Tyr Val Trp Lys Gln Met Gly Ala	
300 305 310 315	
tct acc gaa gga atc act cac tct ttg gcg aac gcg gtt gtg gtt gaa	2091
Ser Thr Glu Gly Ile Thr His Ser Leu Ala Asn Ala Val Val Val Glu	
320 325 330	
gga ggt aag aag ttg ggt tac aag agc aag gaa atc gag cag aac aat	2139
Gly Gly Lys Lys Leu Gly Tyr Lys Ser Lys Glu Ile Glu Gln Asn Asn	
335 340 345	
ggt ggc cat cct gac cac ccc tgt ggt ttc tgt tac ttg ggc tgt aag	2187
Gly Gly His Pro Asp His Pro Cys Gly Phe Cys Tyr Leu Gly Cys Lys	
350 355 360	
tac ggt att aag cag ggt tct gtg aat aac tgg ttt aga gac gca gct	2235
Tyr Gly Ile Lys Gln Gly Ser Val Asn Asn Trp Phe Arg Asp Ala Ala	
365 370 375	
gcc cac ggg tcc aag ttc atg caa caa gtc aga gtt gtg caa atc ctc	2283
Ala His Gly Ser Lys Phe Met Gln Gln Val Arg Val Val Gln Ile Leu	
380 385 390 395	
cac aat aaa ggc gtc gct tat ggc atc ttg tgt gag gat gtc gag acc	2331
His Asn Lys Gly Val Ala Tyr Gly Ile Leu Cys Glu Asp Val Glu Thr	
400 405 410	
gga gtc aaa ttc act atc agt ggc ccc aaa aag ttt gtt gtt tct gca	2379
Gly Val Lys Phe Thr Ile Ser Gly Pro Lys Lys Phe Val Val Ser Ala	
415 420 425	
ggt tct ttg aac acg cca acg gtg ttg acc aac tcc gga ttc aag aac	2427
Gly Ser Leu Asn Thr Pro Thr Val Leu Thr Asn Ser Gly Phe Lys Asn	
430 435 440	
aaa cac atc ggt aag aac ttg acg ttg cac cca gtt tcg acc gtg ttt	2475
Lys His Ile Gly Lys Asn Leu Thr Leu His Pro Val Ser Thr Val Phe	
445 450 455	
ggt gac ttt ggc aga gac gtg caa gcc gac cat ttc cac aaa tct att	2523
Gly Asp Phe Gly Arg Asp Val Gln Ala Asp His Phe His Lys Ser Ile	
460 465 470 475	
atg act tcg ctc tgt tac gaa gtc gct gac ttg gac ggc aag ggc cac	2571
Met Thr Ser Leu Cys Tyr Glu Val Ala Asp Leu Asp Gly Lys Gly His	
480 485 490	
gga tgc aga atc gag acc atc ttg aac gct cca ttc atc caa gct tct	2619
Gly Cys Arg Ile Glu Thr Ile Leu Asn Ala Pro Phe Ile Gln Ala Ser	
495 500 505	
ttg ttg cca tgg aga gga agc gac gag gtc aga aga gac ttg ttg cgt	2667



ES 2 379 776 T3

Leu	Leu	Pro	Trp	Arg	Gly	Ser	Asp	Glu	Val	Arg	Arg	Asp	Leu	Leu	Arg		
		510					515					520					
tac	aac	aac	atg	gtg	gcc	atg	ttg	ctt	atc	acc	cgt	gac	acc	acc	agt		2715
Tyr	Asn	Asn	Met	Val	Ala	Met	Leu	Leu	Ile	Thr	Arg	Asp	Thr	Thr	Ser		
	525					530					535						
ggt	tca	ggt	tct	gct	gac	cca	aag	aag	ccc	gac	gct	ttg	att	gtc	gac		2763
Gly	Ser	Val	Ser	Ala	Asp	Pro	Lys	Lys	Pro	Asp	Ala	Leu	Ile	Val	Asp		
540					545					550					555		
tat	gac	atc	aac	aag	ttt	gac	aag	aat	gcc	atc	ttg	caa	gct	ttc	ttg		2811
Tyr	Asp	Ile	Asn	Lys	Phe	Asp	Lys	Asn	Ala	Ile	Leu	Gln	Ala	Phe	Leu		
			560						565					570			
atc	acc	tcc	gac	atg	ttg	tac	atc	gaa	ggt	gcc	aag	aga	atc	ctc	agt		2859
Ile	Thr	Ser	Asp	Met	Leu	Tyr	Ile	Glu	Gly	Ala	Lys	Arg	Ile	Leu	Ser		
			575					580					585				
cca	cag	gca	tgg	gtg	cca	atc	ttt	gag	tcg	aac	aag	cca	aag	gag	caa		2907
Pro	Gln	Ala	Trp	Val	Pro	Ile	Phe	Glu	Ser	Asn	Lys	Pro	Lys	Glu	Gln		
		590					595					600					
aga	aca	atc	aag	gac	aag	gac	tat	gtc	gaa	tgg	aga	gcc	aag	gct	gcc		2955
Arg	Thr	Ile	Lys	Asp	Lys	Asp	Tyr	Val	Glu	Trp	Arg	Ala	Lys	Ala	Ala		
	605					610						615					
aag	ata	cct	ttc	gac	acc	tac	ggt	tct	gcc	tat	ggg	tcc	gca	cat	caa		3003
Lys	Ile	Pro	Phe	Asp	Thr	Tyr	Gly	Ser	Ala	Tyr	Gly	Ser	Ala	His	Gln		
620					625					630					635		
atg	tcc	acc	tgt	cgt	atg	tcc	gga	aag	ggt	cct	aaa	tac	ggc	gcc	gtt		3051
Met	Ser	Thr	Cys	Arg	Met	Ser	Gly	Lys	Gly	Pro	Lys	Tyr	Gly	Ala	Val		
				640					645					650			
gat	acc	gat	ggt	aga	ttg	ttt	gaa	tgt	tcg	aat	gtc	tat	gtt	gct	gat		3099
Asp	Thr	Asp	Gly	Arg	Leu	Phe	Glu	Cys	Ser	Asn	Val	Tyr	Val	Ala	Asp		
			655				660						665				
gct	agt	ggt	ttg	cct	act	gcc	agc	ggt	gcc	aac	cca	atg	atc	tcc	acc		3147
Ala	Ser	Val	Leu	Pro	Thr	Ala	Ser	Gly	Ala	Asn	Pro	Met	Ile	Ser	Thr		
		670					675					680					
atg	acg	ttt	gct	aga	cag	att	gcg	tta	ggt	ttg	gct	gac	tct	ttg	aag		3195
Met	Thr	Phe	Ala	Arg	Gln	Ile	Ala	Leu	Gly	Leu	Ala	Asp	Ser	Leu	Lys		
	685					690						695					
acc	aaa	ccc	aag	ttg	tag	agagagacag	aaatacagaca	cttatatact									3243
Thr	Lys	Pro	Lys	Leu													
700																	
agatgtatct	tacaatttat	atcttcgatg	atggctttta	ctatctccta	tgttacacta												3303
taatgacatc	accacatctt	ctactactgt	ctccagtatc	ctccttgctg	ttgaccgtat												3363
ccaccagcct	gttggttgaa	ccccgtgaac	tgtggttgct	gttgagcgta	ccccacgtta												3423

ES 2 379 776 T3

gtgaactgcg gttggtgggt aaactgctgt acgggctggt gttgctggtg ctggttgttc 3483  
tgttggtgct gttggtgctg ttggtgctgt tgttggtggt gtcccgttgg ctggttgtac 3543  
aacgacatga tgttctgctt gtttgtctgc tgggcaacca actgtgggtt attcatctgc 3603  
atcaactgct gctggtgctg agggttggtt ggatccaagt actcctgccc gttggcgtcg 3663  
atataagaaa tctgccccgt gactgggtca gtgtactggt atatctgtgg catgccaccc 3723  
gcttgtgcag gcatgccggg tgccaatggc 3753

<210> 6

<211> 704

<212> PRT

<213> *Candida tropicalis*

<400> 6

ES 2 379 776 T3

Met Asn Thr Phe Leu Pro Asp Val Leu Glu Tyr Lys His Val Asp Thr  
 1 5 10 15

Leu Leu Leu Leu Cys Asp Gly Ile Ile His Glu Thr Thr Val Asp Gln  
 20 25 30

Ile Arg Asp Ala Ile Ala Pro Asp Phe Pro Glu Asp Gln Tyr Glu Glu  
 35 40 45

Tyr Leu Lys Thr Phe Thr Lys Pro Ser Glu Thr Pro Gly Phe Arg Glu  
 50 55 60

Ala Val Tyr Asp Thr Ile Asn Ser Thr Pro Thr Glu Ala Val His Met  
 65 70 75 80

Cys Ile Val Leu Thr Thr Ala Leu Asp Ser Arg Ile Leu Ala Pro Thr  
 85 90 95

Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met Thr Leu Lys Glu Arg  
 100 105 110

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

Arg Leu Phe Arg Leu Ile Ser Ser Leu Thr Leu Thr Thr Phe Thr Arg  
 130 135 140

Leu Ala Ser Asp Leu His Leu Arg Ala Ile His Tyr Pro Gly Arg Asp

ES 2 379 776 T3

145					150					155					160
Leu	Arg	Glu	Lys	Ala	Tyr	Glu	Thr	Gln	Val	Val	Asp	Pro	Phe	Arg	Tyr
				165					170					175	
Leu	Phe	Met	Glu	Lys	Pro	Lys	Phe	Asp	Gly	Thr	Glu	Leu	Tyr	Leu	Pro
			180					185					190		
Asp	Ile	Asp	Val	Ile	Ile	Ile	Gly	Ser	Gly	Ala	Gly	Ala	Gly	Val	Met
		195					200					205			
Ala	His	Thr	Leu	Ala	Asn	Asp	Gly	Tyr	Lys	Thr	Leu	Val	Leu	Glu	Lys
	210					215					220				
Gly	Lys	Tyr	Phe	Ser	Asn	Ser	Glu	Leu	Asn	Phe	Asn	Asp	Ala	Asp	Gly
225					230					235					240
Met	Lys	Glu	Leu	Tyr	Gln	Gly	Lys	Cys	Ala	Leu	Thr	Thr	Thr	Asn	Gln
				245					250					255	
Gln	Met	Phe	Ile	Leu	Ala	Gly	Ser	Thr	Leu	Gly	Gly	Gly	Thr	Thr	Val
			260					265					270		
Asn	Trp	Ser	Ala	Cys	Leu	Lys	Thr	Pro	Phe	Lys	Val	Arg	Lys	Glu	Trp
		275					280					285			
Tyr	Asp	Glu	Phe	Gly	Leu	Glu	Phe	Ala	Ala	Asp	Glu	Ala	Tyr	Asp	Lys
	290					295					300				
Ala	Gln	Asp	Tyr	Val	Trp	Lys	Gln	Met	Gly	Ala	Ser	Thr	Glu	Gly	Ile
305					310					315					320
Thr	His	Ser	Leu	Ala	Asn	Ala	Val	Val	Val	Glu	Gly	Gly	Lys	Lys	Leu
				325					330					335	
Gly	Tyr	Lys	Ser	Lys	Glu	Ile	Glu	Gln	Asn	Asn	Gly	Gly	His	Pro	Asp
			340					345					350		
His	Pro	Cys	Gly	Phe	Cys	Tyr	Leu	Gly	Cys	Lys	Tyr	Gly	Ile	Lys	Gln
		355					360					365			
Gly	Ser	Val	Asn	Asn	Trp	Phe	Arg	Asp	Ala	Ala	Ala	His	Gly	Ser	Lys
	370					375					380				

ES 2 379 776 T3

Phe Met Gln Gln Val Arg Val Val Gln Ile Leu His Asn Lys Gly Val  
 385 390 395 400

Ala Tyr Gly Ile Leu Cys Glu Asp Val Glu Thr Gly Val Lys Phe Thr  
 405 410 415

Ile Ser Gly Pro Lys Lys Phe Val Val Ser Ala Gly Ser Leu Asn Thr  
 420 425 430

Pro Thr Val Leu Thr Asn Ser Gly Phe Lys Asn Lys His Ile Gly Lys  
 435 440 445

Asn Leu Thr Leu His Pro Val Ser Thr Val Phe Gly Asp Phe Gly Arg  
 450 455 460

Asp Val Gln Ala Asp His Phe His Lys Ser Ile Met Thr Ser Leu Cys  
 465 470 475 480

Tyr Glu Val Ala Asp Leu Asp Gly Lys Gly His Gly Cys Arg Ile Glu  
 485 490 495

Thr Ile Leu Asn Ala Pro Phe Ile Gln Ala Ser Leu Leu Pro Trp Arg  
 500 505 510

Gly Ser Asp Glu Val Arg Arg Asp Leu Leu Arg Tyr Asn Asn Met Val  
 515 520 525

Ala Met Leu Leu Ile Thr Arg Asp Thr Thr Ser Gly Ser Val Ser Ala  
 530 535 540

Asp Pro Lys Lys Pro Asp Ala Leu Ile Val Asp Tyr Asp Ile Asn Lys  
 545 550 555 560

Phe Asp Lys Asn Ala Ile Leu Gln Ala Phe Leu Ile Thr Ser Asp Met  
 565 570 575

Leu Tyr Ile Glu Gly Ala Lys Arg Ile Leu Ser Pro Gln Ala Trp Val  
 580 585 590

Pro Ile Phe Glu Ser Asn Lys Pro Lys Glu Gln Arg Thr Ile Lys Asp  
 595 600 605

ES 2 379 776 T3

Lys Asp Tyr Val Glu Trp Arg Ala Lys Ala Ala Lys Ile Pro Phe Asp  
610 615 620

Thr Tyr Gly Ser Ala Tyr Gly Ser Ala His Gln Met Ser Thr Cys Arg  
625 630 635 640

Met Ser Gly Lys Gly Pro Lys Tyr Gly Ala Val Asp Thr Asp Gly Arg  
645 650 655

Leu Phe Glu Cys Ser Asn Val Tyr Val Ala Asp Ala Ser Val Leu Pro  
660 665 670

Thr Ala Ser Gly Ala Asn Pro Met Ile Ser Thr Met Thr Phe Ala Arg  
675 680 685

Gln Ile Ala Leu Gly Leu Ala Asp Ser Leu Lys Thr Lys Pro Lys Leu  
690 695 700

<210> 7

<211> 3545

<212> ADN

<213> *Candida tropicalis*

<220>

<221> CDS

<222> (1179)..(3290)

<400> 7

ES 2 379 776 T3

ccttaattaa agtctccaag ttgaccgacg cccaagtcac gtaccacttt atttccggtt	60
acacttccaa gatggctggt actgaagaag gtgtcacgga accacaagct actttctccg	120
cttgtttcgg tcaaccattc ttggtggtgc acccaatgaa gtacgctcaa caattgtctg	180
acaagatctc gcaacacaag gctaacgcct ggttggtgaa caccggttgg gttggttctt	240
ctgctgctag aggtggtaag agatgctcat tgaagtacac cagagccatt ttggacgcta	300
tccactctgg tgaattgtcc aagggtgaat acgaaacttt cccagtcttc aacttgaatg	360
tccaacctc ctgtccaggt gtccaagtg aatcttgaa cccaaccaag gcctggaccg	420
gaaggtggtg actccttcaa caaggaaatc aagtctttgg ctggtaagtt tgctgaaaac	480
ttcaagacct atgctgacca agctaccgct gaagtgagag ctgcaggtcc agaagcttaa	540
agatatttat tcattattta gtttgcctat ttatttctca ttacccatca tcattcaaca	600
ctatatataa agttacttcg gatatcattg taatcgtgcg tgcgcaatt ggatgatttg	660

ES 2 379 776 T3

gaactgCGct tGaaacggat tcatgcacga agcggagata aaagattacg taatttatct 720  
 cctgagacaa ttttagccgt gttcacacgc ctttctttgt tctgagcgaa ggataaataa 780  
 ttagacttcc acagctcatt ctaatttccg tcacgcgaat attgaagggg ggtacatgtg 840  
 gccgctgaat gtggggggcag taaacgcagt ctctcctctc ccaggaatag tgcaacggag 900  
 gaaggataac ggatagaaag cggaatgcga ggaaaatTTT gaacgcgcaa gaaaagcaat 960  
 atccgggcta ccaggTTTTg agccagggaa cacactccta tttctgctca atgactgaac 1020  
 atagaaaaaa caccaagacg caatgaaacg cacatggaca tttagacctc cccacatgtg 1080  
 atagtttgtc ttaacagaaa agtataataa gaacccatgc cgtccctttt ctttcgccgc 1140  
 ttcaactttt ttttttttat cttacacaca tcacgacc atg gct cca ttt ttg ccc 1196  
 Met Ala Pro Phe Leu Pro  
 1 5  
 gac cag gtc gac tac aaa cac gtc gac acc ctt atg tta tta tgt gac 1244  
 Asp Gln Val Asp Tyr Lys His Val Asp Thr Leu Met Leu Leu Cys Asp  
 10 15 20  
 ggg atc atc cac gaa acc acc gtg gac gaa atc aaa gac gtc att gcc 1292  
 Gly Ile Ile His Glu Thr Thr Val Asp Glu Ile Lys Asp Val Ile Ala  
 25 30 35  
 cct gac ttc ccc gcc gac aaa tac gag gag tac gtc agg aca ttc acc 1340  
 Pro Asp Phe Pro Ala Asp Lys Tyr Glu Glu Tyr Val Arg Thr Phe Thr  
 40 45 50  
 aaa ccc tcc gaa acc cca ggg ttc agg gaa acc gtc tac aac acc gtc 1388  
 Lys Pro Ser Glu Thr Pro Gly Phe Arg Glu Thr Val Tyr Asn Thr Val  
 55 60 65 70  
 aac gca aac acc atg gat gca atc cac cag ttc att atc ttg acc aat 1436  
 Asn Ala Asn Thr Met Asp Ala Ile His Gln Phe Ile Ile Leu Thr Asn  
 75 80 85  
 gtt ttg gga tca agg gtc ttg gca cca gct ttg acc aac tcg ttg act 1484  
 Val Leu Gly Ser Arg Val Leu Ala Pro Ala Leu Thr Asn Ser Leu Thr  
 90 95 100  
 cct atc aag gac atg agc ttg gaa gac cgt gaa aag ttg tta gcc tcg 1532  
 Pro Ile Lys Asp Met Ser Leu Glu Asp Arg Glu Lys Leu Leu Ala Ser  
 105 110 115  
 tgg cgt gac tcc cct att gct gct aaa agg aag ttg ttc agg ttg gtt 1580  
 Trp Arg Asp Ser Pro Ile Ala Ala Lys Arg Lys Leu Phe Arg Leu Val  
 120 125 130  
 tct acg ctt acc ttg gtc acg ttc acg aga ttg gcc aat gag ttg cat 1628  
 Ser Thr Leu Thr Leu Val Thr Phe Thr Arg Leu Ala Asn Glu Leu His  
 135 140 145 150



ES 2 379 776 T3

ttg aaa gcc att cat tat cca gga aga gaa gac cgt gaa aag gct tat	1676
Leu Lys Ala Ile His Tyr Pro Gly Arg Glu Asp Arg Glu Lys Ala Tyr	
155 160 165	
gaa acc cag gag att gac cct ttt aag tac cag ttt ttg gaa aaa ccg	1724
Glu Thr Gln Glu Ile Asp Pro Phe Lys Tyr Gln Phe Leu Glu Lys Pro	
170 175 180	
aag ttt tac ggc gct gag ttg tac ttg cca gat att gat gtg atc att	1772
Lys Phe Tyr Gly Ala Glu Leu Tyr Leu Pro Asp Ile Asp Val Ile Ile	
185 190 195	
att gga tct ggg gcc ggt gct ggt gtc gtg gcc cac act ttg acc aac	1820
Ile Gly Ser Gly Ala Gly Ala Gly Val Val Ala His Thr Leu Thr Asn	
200 205 210	
gac ggc ttc aag agt ttg gtt ttg gaa aag ggc aga tac ttt agc aac	1868
Asp Gly Phe Lys Ser Leu Val Leu Glu Lys Gly Arg Tyr Phe Ser Asn	
215 220 225 230	
tcc gag ttg aac ttt gat gac aag gac ggg gtt caa gaa tta tac caa	1916
Ser Glu Leu Asn Phe Asp Asp Lys Asp Gly Val Gln Glu Leu Tyr Gln	
235 240 245	
agt gga ggt act ttg acc acc gtc aac cag cag ttg ttt gtt ctt gct	1964
Ser Gly Gly Thr Leu Thr Thr Val Asn Gln Gln Leu Phe Val Leu Ala	
250 255 260	
ggt tcc act ttt ggt ggt ggt acc act gtc aat tgg tcg gcc tgt ctt	2012
Gly Ser Thr Phe Gly Gly Gly Thr Thr Val Asn Trp Ser Ala Cys Leu	
265 270 275	
aaa acg cca ttc aag gtg cgt aag gaa tgg tat gat gag ttt ggc gtt	2060
Lys Thr Pro Phe Lys Val Arg Lys Glu Trp Tyr Asp Glu Phe Gly Val	
280 285 290	
gac ttt gct gcc gat gaa gcc tac gac aaa gca cag gat tat gtt tgg	2108
Asp Phe Ala Ala Asp Glu Ala Tyr Asp Lys Ala Gln Asp Tyr Val Trp	
295 300 305 310	
cag caa atg gga gct tct acc gaa ggc atc acc cac tct ttg gct aac	2156
Gln Gln Met Gly Ala Ser Thr Glu Gly Ile Thr His Ser Leu Ala Asn	
315 320 325	
gag att att att gaa ggt ggc aag aaa tta ggt tac aag gcc aag gta	2204
Glu Ile Ile Ile Glu Gly Gly Lys Lys Leu Gly Tyr Lys Ala Lys Val	
330 335 340	
tta gac caa aac agc ggt ggt cat cct cat cac aga tgc ggt ttc tgt	2252
Leu Asp Gln Asn Ser Gly Gly His Pro His His Arg Cys Gly Phe Cys	
345 350 355	
tat ttg ggt tgt aag cac ggt atc aag cag ggc tct gtt aat aac tgg	2300
Tyr Leu Gly Cys Lys His Gly Ile Lys Gln Gly Ser Val Asn Asn Trp	
360 365 370	
ttt aga gac gca gct gcc cac ggt tct cag ttc atg caa cag gtt aga	2348

ES 2 379 776 T3

Phe	Arg	Asp	Ala	Ala	Ala	His	Gly	Ser	Gln	Phe	Met	Gln	Gln	Val	Arg	
375					380					385					390	
ggt	ttg	caa	atc	ctt	aac	aag	aag	ggc	atc	gct	tat	ggt	atc	ttg	tgt	2396
Val	Leu	Gln	Ile	Leu	Asn	Lys	Lys	Gly	Ile	Ala	Tyr	Gly	Ile	Leu	Cys	
				395				400						405		
gag	gat	ggt	gta	acc	ggt	gcc	aag	ttc	acc	att	act	ggc	ccc	aaa	aag	2444
Glu	Asp	Val	Val	Thr	Gly	Ala	Lys	Phe	Thr	Ile	Thr	Gly	Pro	Lys	Lys	
			410					415					420			
ttt	ggt	ggt	gcc	gcc	ggc	gcc	tta	aac	act	cca	tct	gtg	ttg	gtc	aac	2492
Phe	Val	Val	Ala	Ala	Gly	Ala	Leu	Asn	Thr	Pro	Ser	Val	Leu	Val	Asn	
		425					430					435				
tcc	gga	ttc	aag	aac	aag	aac	atc	ggt	aag	aac	tta	act	ttg	cat	cca	2540
Ser	Gly	Phe	Lys	Asn	Lys	Asn	Ile	Gly	Lys	Asn	Leu	Thr	Leu	His	Pro	
	440					445				450						
ggt	tct	gtc	gtg	ttt	ggt	gat	ttt	ggc	aaa	gac	ggt	caa	gca	gat	cac	2588
Val	Ser	Val	Val	Phe	Gly	Asp	Phe	Gly	Lys	Asp	Val	Gln	Ala	Asp	His	
455					460					465					470	
ttc	cac	aac	tcc	atc	atg	act	gct	ctt	tgt	tca	gaa	gcc	gct	gat	tta	2636
Phe	His	Asn	Ser	Ile	Met	Thr	Ala	Leu	Cys	Ser	Glu	Ala	Ala	Asp	Leu	
				475					480					485		
gac	ggc	aag	ggt	cat	gga	tgc	aga	att	gaa	acc	atc	ttg	aac	gct	cca	2684
Asp	Gly	Lys	Gly	His	Gly	Cys	Arg	Ile	Glu	Thr	Ile	Leu	Asn	Ala	Pro	
			490					495					500			
ttc	atc	cag	gct	tca	ttc	tta	cca	tgg	aga	ggt	agt	aac	gag	gct	aga	2732
Phe	Ile	Gln	Ala	Ser	Phe	Leu	Pro	Trp	Arg	Gly	Ser	Asn	Glu	Ala	Arg	
		505					510					515				
cga	gac	ttg	ttg	cgt	tac	aac	aac	atg	gtg	gcc	atg	tta	ctt	ctt	agt	2780
Arg	Asp	Leu	Leu	Arg	Tyr	Asn	Asn	Met	Val	Ala	Met	Leu	Leu	Leu	Ser	
	520					525					530					
cgt	gat	acc	acc	agt	ggt	tcc	ggt	tcg	tcc	cat	cca	act	aaa	cct	gaa	2828
Arg	Asp	Thr	Thr	Ser	Gly	Ser	Val	Ser	Ser	His	Pro	Thr	Lys	Pro	Glu	
535					540					545				550		
gca	tta	ggt	gtc	gag	tac	gac	gtg	aac	aag	ttt	gac	aga	aac	tcc	atc	2876
Ala	Leu	Val	Val	Glu	Tyr	Asp	Val	Asn	Lys	Phe	Asp	Arg	Asn	Ser	Ile	
				555					560					565		
ttg	cag	gca	ttg	ttg	gtc	act	gct	gac	ttg	ttg	tac	att	caa	ggt	gcc	2924
Leu	Gln	Ala	Leu	Leu	Val	Thr	Ala	Asp	Leu	Leu	Tyr	Ile	Gln	Gly	Ala	
				570				575					580			
aag	aga	atc	ctt	agt	ccc	caa	cca	tgg	gtg	cca	att	ttt	gaa	tcc	gac	2972
Lys	Arg	Ile	Leu	Ser	Pro	Gln	Pro	Trp	Val	Pro	Ile	Phe	Glu	Ser	Asp	
		585					590					595				
aag	cca	aag	gat	aag	aga	tca	atc	aag	gac	gag	gac	tat	gtc	gaa	tgg	3020
Lys	Pro	Lys	Asp	Lys	Arg	Ser	Ile	Lys	Asp	Glu	Asp	Tyr	Val	Glu	Trp	

ES 2 379 776 T3

600	605	610	
aga gcc aag gtt gcc aag att cct ttt gac acc tac ggc tcg cct tat			3068
Arg Ala Lys Val Ala Lys Ile Pro Phe Asp Thr Tyr Gly Ser Pro Tyr			
615	620	625	630
ggt tcg gcg cat caa atg tct tct tgt cgt atg tca ggt aag ggt cct			3116
Gly Ser Ala His Gln Met Ser Ser Cys Arg Met Ser Gly Lys Gly Pro			
	635	640	645
aaa tac ggt gct gtt gat acc gat ggt aga ttg ttt gaa tgt tcg aat			3164
Lys Tyr Gly Ala Val Asp Thr Asp Gly Arg Leu Phe Glu Cys Ser Asn			
	650	655	660
gtt tat gtt gct gac gct agt ctt ttg cca act gct agc ggt gct aat			3212
Val Tyr Val Ala Asp Ala Ser Leu Leu Pro Thr Ala Ser Gly Ala Asn			
	665	670	675
cct atg gtc acc acc atg act ctt gca aga cat gtt gcg tta ggt ttg			3260
Pro Met Val Thr Thr Met Thr Leu Ala Arg His Val Ala Leu Gly Leu			
	680	685	690
gca gac tcc ttg aag acc aag gcc aag ttg tagttctgta tacgtatctt			3310
Ala Asp Ser Leu Lys Thr Lys Ala Lys Leu			
695	700		
ataatttaga tttcctttta ttgacggtaa acattcagga taggtactac ccttgctgca			3370
aaagcccagc acgcccgaat cgcgatgact tgagcgaagc aaacacgcac acaaaagggg			3430
tacacaaaaa ataacgagat gcccttgaag cacacacca aacacgatgg aacaacaagat			3490
ggccctagaa agtacaaaaa aagtaaaggc acttgattcc gccattaat taagg			3545

<210> 8

<211> 704

<212> PRT

<213> *Candida tropicalis*

<400> 8

ES 2 379 776 T3

Met Ala Pro Phe Leu Pro Asp Gln Val Asp Tyr Lys His Val Asp Thr  
1                    5                    10                    15

Leu Met Leu Leu Cys Asp Gly Ile Ile His Glu Thr Thr Val Asp Glu  
          20                    25                    30

Ile Lys Asp Val Ile Ala Pro Asp Phe Pro Ala Asp Lys Tyr Glu Glu  
          35                    40                    45

Tyr Val Arg Thr Phe Thr Lys Pro Ser Glu Thr Pro Gly Phe Arg Glu  
          50                    55                    60

ES 2 379 776 T3

Thr Val Tyr Asn Thr Val Asn Ala Asn Thr Met Asp Ala Ile His Gln  
65 70 75 80

Phe Ile Ile Leu Thr Asn Val Leu Gly Ser Arg Val Leu Ala Pro Ala  
85 90 95

Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met Ser Leu Glu Asp Arg  
100 105 110

Glu Lys Leu Leu Ala Ser Trp Arg Asp Ser Pro Ile Ala Ala Lys Arg  
115 120 125

Lys Leu Phe Arg Leu Val Ser Thr Leu Thr Leu Val Thr Phe Thr Arg  
130 135 140

Leu Ala Asn Glu Leu His Leu Lys Ala Ile His Tyr Pro Gly Arg Glu  
145 150 155 160

Asp Arg Glu Lys Ala Tyr Glu Thr Gln Glu Ile Asp Pro Phe Lys Tyr  
165 170 175

Gln Phe Leu Glu Lys Pro Lys Phe Tyr Gly Ala Glu Leu Tyr Leu Pro  
180 185 190

Asp Ile Asp Val Ile Ile Ile Gly Ser Gly Ala Gly Ala Gly Val Val  
195 200 205

Ala His Thr Leu Thr Asn Asp Gly Phe Lys Ser Leu Val Leu Glu Lys  
210 215 220

Gly Arg Tyr Phe Ser Asn Ser Glu Leu Asn Phe Asp Asp Lys Asp Gly  
225 230 235 240

Val Gln Glu Leu Tyr Gln Ser Gly Gly Thr Leu Thr Thr Val Asn Gln  
245 250 255

Gln Leu Phe Val Leu Ala Gly Ser Thr Phe Gly Gly Gly Thr Thr Val  
260 265 270

Asn Trp Ser Ala Cys Leu Lys Thr Pro Phe Lys Val Arg Lys Glu Trp  
275 280 285

ES 2 379 776 T3

Tyr Asp Glu Phe Gly Val Asp Phe Ala Ala Asp Glu Ala Tyr Asp Lys  
 290 295 300

Ala Gln Asp Tyr Val Trp Gln Gln Met Gly Ala Ser Thr Glu Gly Ile  
 305 310 315 320

Thr His Ser Leu Ala Asn Glu Ile Ile Ile Glu Gly Gly Lys Lys Leu  
 325 330 335

Gly Tyr Lys Ala Lys Val Leu Asp Gln Asn Ser Gly Gly His Pro His  
 340 345 350

His Arg Cys Gly Phe Cys Tyr Leu Gly Cys Lys His Gly Ile Lys Gln  
 355 360 365

Gly Ser Val Asn Asn Trp Phe Arg Asp Ala Ala Ala His Gly Ser Gln  
 370 375 380

Phe Met Gln Gln Val Arg Val Leu Gln Ile Leu Asn Lys Lys Gly Ile  
 385 390 395 400

Ala Tyr Gly Ile Leu Cys Glu Asp Val Val Thr Gly Ala Lys Phe Thr  
 405 410 415

Ile Thr Gly Pro Lys Lys Phe Val Val Ala Ala Gly Ala Leu Asn Thr  
 420 425 430

Pro Ser Val Leu Val Asn Ser Gly Phe Lys Asn Lys Asn Ile Gly Lys  
 435 440 445

Asn Leu Thr Leu His Pro Val Ser Val Val Phe Gly Asp Phe Gly Lys  
 450 455 460

Asp Val Gln Ala Asp His Phe His Asn Ser Ile Met Thr Ala Leu Cys  
 465 470 475 480

Ser Glu Ala Ala Asp Leu Asp Gly Lys Gly His Gly Cys Arg Ile Glu  
 485 490 495

Thr Ile Leu Asn Ala Pro Phe Ile Gln Ala Ser Phe Leu Pro Trp Arg  
 500 505 510

Gly Ser Asn Glu Ala Arg Arg Asp Leu Leu Arg Tyr Asn Asn Met Val

ES 2 379 776 T3

	515					520						525			
Ala	Met	Leu	Leu	Leu	Ser	Arg	Asp	Thr	Thr	Ser	Gly	Ser	Val	Ser	Ser
	530					535					540				
His	Pro	Thr	Lys	Pro	Glu	Ala	Leu	Val	Val	Glu	Tyr	Asp	Val	Asn	Lys
545					550					555					560
Phe	Asp	Arg	Asn	Ser	Ile	Leu	Gln	Ala	Leu	Leu	Val	Thr	Ala	Asp	Leu
				565					570					575	
Leu	Tyr	Ile	Gln	Gly	Ala	Lys	Arg	Ile	Leu	Ser	Pro	Gln	Pro	Trp	Val
			580					585					590		
Pro	Ile	Phe	Glu	Ser	Asp	Lys	Pro	Lys	Asp	Lys	Arg	Ser	Ile	Lys	Asp
		595					600					605			
Glu	Asp	Tyr	Val	Glu	Trp	Arg	Ala	Lys	Val	Ala	Lys	Ile	Pro	Phe	Asp
	610					615					620				
Thr	Tyr	Gly	Ser	Pro	Tyr	Gly	Ser	Ala	His	Gln	Met	Ser	Ser	Cys	Arg
625					630					635					640
Met	Ser	Gly	Lys	Gly	Pro	Lys	Tyr	Gly	Ala	Val	Asp	Thr	Asp	Gly	Arg
				645					650					655	
Leu	Phe	Glu	Cys	Ser	Asn	Val	Tyr	Val	Ala	Asp	Ala	Ser	Leu	Leu	Pro
			660					665					670		
Thr	Ala	Ser	Gly	Ala	Asn	Pro	Met	Val	Thr	Thr	Met	Thr	Leu	Ala	Arg
		675					680					685			
His	Val	Ala	Leu	Gly	Leu	Ala	Asp	Ser	Leu	Lys	Thr	Lys	Ala	Lys	Leu
	690					695					700				

<210> 9

<211> 4158

<212> ADN

<213> *Candida tropicalis*

<220>

<221> CDS

<222> (1521)..(3632)

<400> 9



ES 2 379 776 T3

tctcaccaag tacgagaacg agcttggtga tagagcttag acttgtcttt tgtatttgta	60
atctgacggt gaccgtttga gtttttcctg tgatatcaag taaatctggc aaccagcttt	120
ctathttttt tgcaacactt tttctcttca ccactctcag aaccaatgoc accgaagaag	180
ggtctcaacc aggaggaaaa gctctcgtca atcctcgcct ggttccaaag cagccactgc	240
ttctatacgc tcaaagaggt tgaacagaag gctagcaaag cgtgcaagat ctcgtctatg	300
cagatcaagg acttggttgc aaccttagtc aacgaagggt tagtggaaca ggaaaaatgt	360
gggaccacca acttggtactg gtcggtcccg tactcggaac acaaacgcaa gctacagaga	420
tacgagcagc taagacaatc cgttgccaaa cttcaagcga ataaaggcaa gttggcggaa	480
gagttgcaaa acgcgtgtgg tgagcgtgac atggacagca ataggctaaa ccggatgcaa	540
cagtgcgatc agcttgttca cgaggcggca cgcctccagg aggaactaaa actgtcgagg	600
cagagagata ccattgacga gttggttcag gccattgact tcttcaacga gctgatagag	660
accgtcctca gctacatcag ccatcagtcg gggaccagcg tgcggtatt gaaaacggag	720
ttgagatac ccgcagaact agaagaggct cccagataa acaatgccgg agttagtgcg	780
taaactgagc atgcatacgt tggagagaaa tagagaaaca gatttccggt gaaacgctac	840
aacacagacg aggaatacag aatggaacat gacggaaata taatatccga ggaaagacga	900
aagtacgaca tggaaactccg ttactgcaac atcgatcgtg ctagatacga catagaacaa	960
tgttgctatt acatggaaag ctggtgctac aatccagaat acggttgtag tcaagggaga	1020
tgaggctggg agccgagtgg tacataaata ggcatatagg accgtcactt ggtctaggat	1080
cgtgtagagg gtggaagagg taggcaagat ccattctaact ctactgagtg acggctaata	1140
tacgatcagc gttctcaggc gagcacagtc attcctcatt tctgtacata cgttgcccct	1200
ttatgttttt tttcacagga tgctcacgcc caacatttcc cccacattt tattaccac	1260
attgagccgt caaatgcatt ttttttatcc gtcgcttgct aagacaaaat tccacatgct	1320
ttgtctcaga gtatataaac aacggggcaa aaaaacatgg ggттаatagc ttattcgtgg	1380
attgatattt ttatatttta gttcgcccct ttcgccacca agctcaattg gactatttgt	1440
cagtggtgta taagctagag attactagac tgcttttctg attcttgatt ttccttttc	1500
attagtcca gtacctagag atg aat acc ttc ttg cca gac gtg ctc gaa tac	1553
Met Asn Thr Phe Leu Pro Asp Val Leu Glu Tyr	
1 5 10	
aaa cac gtc gac acc ctt ttg tta ttg tgt gac ggg atc atc cac gaa	1601
Lys His Val Asp Thr Leu Leu Leu Leu Cys Asp Gly Ile Ile His Glu	

ES 2 379 776 T3

	15	20	25		
acc aca gtc gat cag atc aag gac gcc att gct ccc gac ttc cct gag				1649	
Thr Thr Val Asp Gln Ile Lys Asp Ala Ile Ala Pro Asp Phe Pro Glu	30	35	40		
gac cag tac gag gag tat ctc aag acc ttc acc aag cca tct gag acc				1697	
Asp Gln Tyr Glu Glu Tyr Leu Lys Thr Phe Thr Lys Pro Ser Glu Thr	45	50	55		
cct ggg ttc aga gaa gcc gtc tac gac acg atc aac gcc acc cca acc				1745	
Pro Gly Phe Arg Glu Ala Val Tyr Asp Thr Ile Asn Ala Thr Pro Thr		65	70	75	
gat gcc gtg cac atg tgt att gtc ttg acc acc gca ttg gac tcc aga				1793	
Asp Ala Val His Met Cys Ile Val Leu Thr Thr Ala Leu Asp Ser Arg		80	85	90	
atc ttg gcc ccc acg ttg acc aac tcg ttg acg cct atc aag gat atg				1841	
Ile Leu Ala Pro Thr Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met		95	100	105	
acc ttg aag gag cgt gaa caa ttg ttg gcc tct tgg cgt gat tcc ccg				1889	
Thr Leu Lys Glu Arg Glu Gln Leu Leu Ala Ser Trp Arg Asp Ser Pro		110	115	120	
att gcg gca aag aga aga ttg ttc aga ttg att tcc tcg ctt acc ttg				1937	
Ile Ala Ala Lys Arg Arg Leu Phe Arg Leu Ile Ser Ser Leu Thr Leu		125	130	135	
acg acg ttt acg aga ttg gcc agc gaa ttg cac ttg aaa gcc atc cac				1985	
Thr Thr Phe Thr Arg Leu Ala Ser Glu Leu His Leu Lys Ala Ile His		140	145	150	155
tac cct ggc aga gac ttg cgt gaa aag gcg tat gaa acc cag gtg gtt				2033	
Tyr Pro Gly Arg Asp Leu Arg Glu Lys Ala Tyr Glu Thr Gln Val Val		160	165	170	
gac cct ttc agg tac tcg ttt atg gag aaa cca aag ttt gac ggc gcc				2081	
Asp Pro Phe Arg Tyr Ser Phe Met Glu Lys Pro Lys Phe Asp Gly Ala		175	180	185	
gaa ttg tac ttg cca gat atc gac gtc atc atc att gga tca ggc gcc				2129	
Glu Leu Tyr Leu Pro Asp Ile Asp Val Ile Ile Ile Gly Ser Gly Ala		190	195	200	
ggt gct ggt gtc atg gcc cac act ctc gcc aac gac ggg ttc aag acc				2177	
Gly Ala Gly Val Met Ala His Thr Leu Ala Asn Asp Gly Phe Lys Thr		205	210	215	
ttg gtt ttg gaa aag gga aag tat ttc agc aac tcc gag ttg aac ttt				2225	
Leu Val Leu Glu Lys Gly Lys Tyr Phe Ser Asn Ser Glu Leu Asn Phe		220	225	230	235
aat gac gct gat ggc gtg aaa gag ttg tac caa ggt aaa ggt gct ttg				2273	
Asn Asp Ala Asp Gly Val Lys Glu Leu Tyr Gln Gly Lys Gly Ala Leu		240	245	250	

ES 2 379 776 T3

gcc acc acc aat cag cag atg ttt att ctt gcc ggt tcc act ttg ggc Ala Thr Thr Asn Gln Gln Met Phe Ile Leu Ala Gly Ser Thr Leu Gly	2321
255 260 265	
ggt ggt acc act gtc aac tgg tct gct tgc ctt aaa aca cca ttt aaa Gly Gly Thr Thr Val Asn Trp Ser Ala Cys Leu Lys Thr Pro Phe Lys	2369
270 275 280	
gtg cgt aag gag tgg tac gac gag ttt ggt ctt gaa ttt gct gcc gat Val Arg Lys Glu Trp Tyr Asp Glu Phe Gly Leu Glu Phe Ala Ala Asp	2417
285 290 295	
gaa gcc tac gac aaa gcg cag gat tat gtt tgg aaa caa atg ggt gct Glu Ala Tyr Asp Lys Ala Gln Asp Tyr Val Trp Lys Gln Met Gly Ala	2465
300 305 310 315	
tca aca gat gga atc act cac tcc ttg gcc aac gaa gtt gtg gtt gaa Ser Thr Asp Gly Ile Thr His Ser Leu Ala Asn Glu Val Val Val Glu	2513
320 325 330	
gga ggt aag aag ttg ggc tac aag agc aag gaa att gag cag aac aac Gly Gly Lys Lys Leu Gly Tyr Lys Ser Lys Glu Ile Glu Gln Asn Asn	2561
335 340 345	
ggt ggc cac cct gac cac cca tgt ggt ttc tgt tac ttg ggc tgt aag Gly Gly His Pro Asp His Pro Cys Gly Phe Cys Tyr Leu Gly Cys Lys	2609
350 355 360	
tac ggt att aaa cag ggt tct gtg aat aac tgg ttt aga gac gca gct Tyr Gly Ile Lys Gln Gly Ser Val Asn Asn Trp Phe Arg Asp Ala Ala	2657
365 370 375	
gcc cac ggg tcc aag ttc atg caa caa gtc aga gtt gtg caa atc ctc Ala His Gly Ser Lys Phe Met Gln Gln Val Arg Val Val Gln Ile Leu	2705
380 385 390 395	
aac aag aat ggc gtc gct tat ggt atc ttg tgt gag gat gtc gaa acc Asn Lys Asn Gly Val Ala Tyr Gly Ile Leu Cys Glu Asp Val Glu Thr	2753
400 405 410	
gga gtc agg ttc act att agt ggc ccc aaa aag ttt gtt gtt tct gct Gly Val Arg Phe Thr Ile Ser Gly Pro Lys Lys Phe Val Val Ser Ala	2801
415 420 425	
ggt tct ttg aac acg cca act gtg ttg acc aac tcc gga ttc aag aac Gly Ser Leu Asn Thr Pro Thr Val Leu Thr Asn Ser Gly Phe Lys Asn	2849
430 435 440	
aag cac att ggt aag aac ttg acg ttg cac cca gtt tcc acc gtg ttt Lys His Ile Gly Lys Asn Leu Thr Leu His Pro Val Ser Thr Val Phe	2897
445 450 455	
ggt gac ttt ggc aga gac gtg caa gcc gac cat ttc cac aaa tct att Gly Asp Phe Gly Arg Asp Val Gln Ala Asp His Phe His Lys Ser Ile	2945
460 465 470 475	

ES 2 379 776 T3

atg act tcg ctt tgt tac gag gtt gct gac ttg gac ggc aag ggc cac	2993
Met Thr Ser Leu Cys Tyr Glu Val Ala Asp Leu Asp Gly Lys Gly His	
480 485 490	
gga tgc aga atc gaa acc atc ttg aac gct cca ttc atc caa gct tct	3041
Gly Cys Arg Ile Glu Thr Ile Leu Asn Ala Pro Phe Ile Gln Ala Ser	
495 500 505	
ttg ttg cca tgg aga gga agt gac gag gtc aga aga gac ttg ttg cgt	3089
Leu Leu Pro Trp Arg Gly Ser Asp Glu Val Arg Arg Asp Leu Leu Arg	
510 515 520	
tac aac aac atg gtg gcc atg ttg ctt atc acg cgt gat acc acc agt	3137
Tyr Asn Asn Met Val Ala Met Leu Leu Ile Thr Arg Asp Thr Thr Ser	
525 530 535	
ggt tca gtt tct gct gac cca aag aag ccc gac gct ttg att gtc gac	3185
Gly Ser Val Ser Ala Asp Pro Lys Lys Pro Asp Ala Leu Ile Val Asp	
540 545 550 555	
tat gag att aac aag ttt gac aag aat gcc atc ttg caa gct ttc ttg	3233
Tyr Glu Ile Asn Lys Phe Asp Lys Asn Ala Ile Leu Gln Ala Phe Leu	
560 565 570	
atc act tcc gac atg ttg tac att gaa ggt gcc aag aga atc ctc agt	3281
Ile Thr Ser Asp Met Leu Tyr Ile Glu Gly Ala Lys Arg Ile Leu Ser	
575 580 585	
cca cag cca tgg gtg cca atc ttt gag tcg aac aag cca aag gag caa	3329
Pro Gln Pro Trp Val Pro Ile Phe Glu Ser Asn Lys Pro Lys Glu Gln	
590 595 600	
aga acg atc aag gac aag gac tat gtt gag tgg aga gcc aag gct gct	3377
Arg Thr Ile Lys Asp Lys Asp Tyr Val Glu Trp Arg Ala Lys Ala Ala	
605 610 615	
aag ata cct ttc gac acc tac ggt tct gca tat ggg tcc gca cat caa	3425
Lys Ile Pro Phe Asp Thr Tyr Gly Ser Ala Tyr Gly Ser Ala His Gln	
620 625 630 635	
atg tcc acc tgt cgt atg tcc gga aag ggt cct aaa tac ggt gct gtt	3473
Met Ser Thr Cys Arg Met Ser Gly Lys Gly Pro Lys Tyr Gly Ala Val	
640 645 650	
gat act gat ggt aga ttg ttt gaa tgt tcg aat gtc tat gtt gct gat	3521
Asp Thr Asp Gly Arg Leu Phe Glu Cys Ser Asn Val Tyr Val Ala Asp	
655 660 665	
gct agt gtt ttg cct act gcc agc ggt gcc aac cca atg ata tcc acc	3569
Ala Ser Val Leu Pro Thr Ala Ser Gly Ala Asn Pro Met Ile Ser Thr	
670 675 680	
atg acc ttt gct aga cag att gcg tta ggt ttg gct gac tcc ttg aag	3617
Met Thr Phe Ala Arg Gln Ile Ala Leu Gly Leu Ala Asp Ser Leu Lys	
685 690 695	
acc aaa ccc aag ttg tagagagacg gaaatacgac acttatatac tagatgtatc	3672

ES 2 379 776 T3

Thr Lys Pro Lys Leu  
700

ttacaattta tattctcgat gatggctttt actatctcct atggttacact ataatgacat	3732
caccacaacc tctactactg tctccagtat cctccttgct gttgaccgta cccaccagcc	3792
tgttgattga accctgtgaa ctgtgggtgc tgttgagcgt accccacggt agtgaactgc	3852
ggttggtggg caaactgctg tacgggctgt tgctgctgct gctggtggtg ttggtggtgt	3912
tgccccgtgg gctggttgta caacgacatg atgttctgct tgtttgtctg ttgggcaacc	3972
aactgtgggt tattcatctg catcaactgc tgctggtggt gagggttggt tggatccaag	4032
tactcttgcc cgttggcgtc gatataagaa atctgccccg tgactgggtc agtgtactgg	4092
tatatctgtg gcatgccacc agcttgtgca ggcatgccgg ttgccaatgg cacctgtgct	4152
tgcgtc	4158

<210> 10

<211> 704

<212> PRT

<213> *Candida tropicalis*

<400> 10

ES 2 379 776 T3

Met Asn Thr Phe Leu Pro Asp Val Leu Glu Tyr Lys His Val Asp Thr  
 1 5 10 15

Leu Leu Leu Leu Cys Asp Gly Ile Ile His Glu Thr Thr Val Asp Gln  
 20 25 30

Ile Lys Asp Ala Ile Ala Pro Asp Phe Pro Glu Asp Gln Tyr Glu Glu  
 35 40 45

Tyr Leu Lys Thr Phe Thr Lys Pro Ser Glu Thr Pro Gly Phe Arg Glu  
 50 55 60

Ala Val Tyr Asp Thr Ile Asn Ala Thr Pro Thr Asp Ala Val His Met  
 65 70 75 80

Cys Ile Val Leu Thr Thr Ala Leu Asp Ser Arg Ile Leu Ala Pro Thr  
 85 90 95

Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met Thr Leu Lys Glu Arg  
 100 105 110

ES 2 379 776 T3

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

Arg Leu Phe Arg Leu Ile Ser Ser Leu Thr Leu Thr Thr Phe Thr Arg  
 130 135 140

Leu Ala Ser Glu Leu His Leu Lys Ala Ile His Tyr Pro Gly Arg Asp  
 145 150 155 160

Leu Arg Glu Lys Ala Tyr Glu Thr Gln Val Val Asp Pro Phe Arg Tyr  
 165 170 175

Ser Phe Met Glu Lys Pro Lys Phe Asp Gly Ala Glu Leu Tyr Leu Pro  
 180 185 190

Asp Ile Asp Val Ile Ile Ile Gly Ser Gly Ala Gly Ala Gly Val Met  
 195 200 205

Ala His Thr Leu Ala Asn Asp Gly Phe Lys Thr Leu Val Leu Glu Lys  
 210 215 220

Gly Lys Tyr Phe Ser Asn Ser Glu Leu Asn Phe Asn Asp Ala Asp Gly  
 225 230 235 240

Val Lys Glu Leu Tyr Gln Gly Lys Gly Ala Leu Ala Thr Thr Asn Gln  
 245 250 255

Gln Met Phe Ile Leu Ala Gly Ser Thr Leu Gly Gly Gly Thr Thr Val  
 260 265 270

Asn Trp Ser Ala Cys Leu Lys Thr Pro Phe Lys Val Arg Lys Glu Trp  
 275 280 285

Tyr Asp Glu Phe Gly Leu Glu Phe Ala Ala Asp Glu Ala Tyr Asp Lys  
 290 295 300

Ala Gln Asp Tyr Val Trp Lys Gln Met Gly Ala Ser Thr Asp Gly Ile  
 305 310 315 320

Thr His Ser Leu Ala Asn Glu Val Val Val Glu Gly Gly Lys Lys Leu  
 325 330 335

Gly Tyr Lys Ser Lys Glu Ile Glu Gln Asn Asn Gly Gly His Pro Asp

ES 2 379 776 T3

	340		345		350														
His	Pro	Cys	Gly	Phe	Cys	Tyr	Leu	Gly	Cys	Lys	Tyr	Gly	Ile	Lys	Gln				
		355					360					365							
Gly	Ser	Val	Asn	Asn	Trp	Phe	Arg	Asp	Ala	Ala	Ala	His	Gly	Ser	Lys				
	370					375						380							
Phe	Met	Gln	Gln	Val	Arg	Val	Val	Gln	Ile	Leu	Asn	Lys	Asn	Gly	Val				
385					390					395									400
Ala	Tyr	Gly	Ile	Leu	Cys	Glu	Asp	Val	Glu	Thr	Gly	Val	Arg	Phe	Thr				
				405					410					415					
Ile	Ser	Gly	Pro	Lys	Lys	Phe	Val	Val	Ser	Ala	Gly	Ser	Leu	Asn	Thr				
			420					425						430					
Pro	Thr	Val	Leu	Thr	Asn	Ser	Gly	Phe	Lys	Asn	Lys	His	Ile	Gly	Lys				
		435					440						445						
Asn	Leu	Thr	Leu	His	Pro	Val	Ser	Thr	Val	Phe	Gly	Asp	Phe	Gly	Arg				
	450					455					460								
Asp	Val	Gln	Ala	Asp	His	Phe	His	Lys	Ser	Ile	Met	Thr	Ser	Leu	Cys				
465					470					475					480				
Tyr	Glu	Val	Ala	Asp	Leu	Asp	Gly	Lys	Gly	His	Gly	Cys	Arg	Ile	Glu				
				485					490					495					
Thr	Ile	Leu	Asn	Ala	Pro	Phe	Ile	Gln	Ala	Ser	Leu	Leu	Pro	Trp	Arg				
			500					505					510						
Gly	Ser	Asp	Glu	Val	Arg	Arg	Asp	Leu	Leu	Arg	Tyr	Asn	Asn	Met	Val				
		515					520					525							
Ala	Met	Leu	Leu	Ile	Thr	Arg	Asp	Thr	Thr	Ser	Gly	Ser	Val	Ser	Ala				
	530					535					540								
Asp	Pro	Lys	Lys	Pro	Asp	Ala	Leu	Ile	Val	Asp	Tyr	Glu	Ile	Asn	Lys				
545					550					555					560				
Phe	Asp	Lys	Asn	Ala	Ile	Leu	Gln	Ala	Phe	Leu	Ile	Thr	Ser	Asp	Met				
				565					570					575					



ES 2 379 776 T3

Leu Tyr Ile Glu Gly Ala Lys Arg Ile Leu Ser Pro Gln Pro Trp Val  
 580 585 590

Pro Ile Phe Glu Ser Asn Lys Pro Lys Glu Gln Arg Thr Ile Lys Asp  
 595 600 605

Lys Asp Tyr Val Glu Trp Arg Ala Lys Ala Ala Lys Ile Pro Phe Asp  
 610 615 620

Thr Tyr Gly Ser Ala Tyr Gly Ser Ala His Gln Met Ser Thr Cys Arg  
 625 630 635 640

Met Ser Gly Lys Gly Pro Lys Tyr Gly Ala Val Asp Thr Asp Gly Arg  
 645 650 655

Leu Phe Glu Cys Ser Asn Val Tyr Val Ala Asp Ala Ser Val Leu Pro  
 660 665 670

Thr Ala Ser Gly Ala Asn Pro Met Ile Ser Thr Met Thr Phe Ala Arg  
 675 680 685

Gln Ile Ala Leu Gly Leu Ala Asp Ser Leu Lys Thr Lys Pro Lys Leu  
 690 695 700

<210> 11

<211> 3753

<212> ADN

<213> *Candida tropicalis*

<220>

<221> CDS

<222> (1099)..(3213)

<400> 11

ES 2 379 776 T3

tgttgataga gcttagactt gtgttttgta tttgtaatct gacgttgatc gtttgatatt 60  
ttcctgtgat atcacgtaaa ttcggcaacc aactttttac tttttgcaac actttctctt 120  
caccactctc agaaccaatg ccaccgaaga agggctctcag ccaggaggaa aagctctcgg 180  
cactcctcac ctggttccaa gccagtcatt gcttctacac actcaaggag gttgaacaga 240  
aggcgagcaa agcgtgcaag atctcgtcta tgcagatcaa ggacttggtt gcaagcttag 300  
tcaacgaagg tttggtagaa caggaaaagt gtgggaccac aaacttgtagc tggtcggtcc 360

ES 2 379 776 T3

agtactcgga attcaaacgg aagctacaga gatacgggca gctaagacaa tcagccgcca	420
aacttcaagc ggataaaggc aagttggcgg aagagttgcg aaacgcatgt ggtgaacggg	480
acatggacaa caataggcaa gaccggatgc aacaatacga tcaccttggt aacgaggcgg	540
cacgtctcca ggaggaacta aaactgtcaa ggcagataga taccattgac gagttagttc	600
aggccattga tttcttcaac gagetgatag agaccgtcct cagctacatc agccatcagt	660
cagggaccag cgtgtcgata ttgaaaacgg agtttgagat acccgcagaa ctagaagagg	720
ccccccagat aagcaatgcc ggagttagt cgtaaatega gcaggcatac attgccctt	780
tgtatTTTTT cacaggatgc tcaccccacc acgcccacaa tttccccca cattttatta	840
cccacattga gccgtcaaat gcattTTTTT atccgtcgct agctaaacca aaattccaca	900
tgcgttgccct cagagtatat aaacaacggg gcaaaaaaca tgggattaat agcttatttg	960
tggattgata tttttatatt ttagttcgcc ctttctacga ccaagctcaa ttggactatt	1020
tgtcagtggg gtataagcta gagattacta gactgctttt ctgattcttg atcatcccct	1080
tagttccagt gcctagag atg aat acc ttc ttg cca gac gtg ctc gaa tac	1131
Met Asn Thr Phe Leu Pro Asp Val Leu Glu Tyr	
1 5 10	
aaa cac gtc gat acc ctt ttg tta tta tgt gac ggg atc atc cac gaa	1179
Lys His Val Asp Thr Leu Leu Leu Leu Cys Asp Gly Ile Ile His Glu	
15 20 25	
acc aca gtc gac cag atc agg gac gcc att gct ccc gac ttc cct gaa	1227
Thr Thr Val Asp Gln Ile Arg Asp Ala Ile Ala Pro Asp Phe Pro Glu	
30 35 40	
gac cag tac gag gag tat ctc aag acc ttc acc aag cca tct gag acc	1275
Asp Gln Tyr Glu Glu Tyr Leu Lys Thr Phe Thr Lys Pro Ser Glu Thr	
45 50 55	
cct ggg ttc aga gaa gcc gtc tac gac acg atc aac agc acc cca acc	1323
Pro Gly Phe Arg Glu Ala Val Tyr Asp Thr Ile Asn Ser Thr Pro Thr	
60 65 70 75	
gag gct gtg cac atg tgt att gta ttg acc acc gca ttg gac tcg aga	1371
Glu Ala Val His Met Cys Ile Val Leu Thr Thr Ala Leu Asp Ser Arg	
80 85 90	
atc ttg gcc ccc acg ttg acc aac tcg ttg acg cct atc aag gat atg	1419
Ile Leu Ala Pro Thr Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met	
95 100 105	
acc ttg aaa gag cgt gaa caa ttg ttg gct gcc tgg cgt gat tcc ccg	1467
Thr Leu Lys Glu Arg Glu Gln Leu Leu Ala Ala Trp Arg Asp Ser Pro	
110 115 120	
atc gcg gcc aag aga aga ttg ttc aga ttg att tcc tca ctt acc ttg	1515

ES 2 379 776 T3

Ile	Ala	Ala	Lys	Arg	Arg	Leu	Phe	Arg	Leu	Ile	Ser	Ser	Leu	Thr	Leu		
	125					130					135						
acg	acc	ttt	acg	aga	ttg	gcc	agc	gac	ttg	cac	ttg	aga	gcc	atc	cac		1563
Thr	Thr	Phe	Thr	Arg	Leu	Ala	Ser	Asp	Leu	His	Leu	Arg	Ala	Ile	His		
140					145					150					155		
tac	cct	ggc	aga	gac	ttg	cgt	gaa	aag	gca	tat	gaa	acc	cag	gtg	gtt		1611
Tyr	Pro	Gly	Arg	Asp	Leu	Arg	Glu	Lys	Ala	Tyr	Glu	Thr	Gln	Val	Val		
				160					165					170			
gac	cct	ttc	agg	tac	tcg	ttt	atg	gaa	aaa	cca	aag	ttt	gac	ggc	acc		1659
Asp	Pro	Phe	Arg	Tyr	Ser	Phe	Met	Glu	Lys	Pro	Lys	Phe	Asp	Gly	Thr		
			175					180					185				
gag	ttg	tac	ttg	cca	gat	atc	gac	gtc	atc	atc	att	gga	tcc	ggt	gcc		1707
Glu	Leu	Tyr	Leu	Pro	Asp	Ile	Asp	Val	Ile	Ile	Ile	Gly	Ser	Gly	Ala		
		190					195					200					
ggt	gct	ggt	gtc	atg	gcc	cac	act	tta	gcc	aac	gac	ggg	tac	aag	acc		1755
Gly	Ala	Gly	Val	Met	Ala	His	Thr	Leu	Ala	Asn	Asp	Gly	Tyr	Lys	Thr		
	205					210					215						
ttg	gtt	ttg	gaa	aag	gga	aag	tat	ttc	agc	aac	tcc	gag	ttg	aac	ttt		1803
Leu	Val	Leu	Glu	Lys	Gly	Lys	Tyr	Phe	Ser	Asn	Ser	Glu	Leu	Asn	Phe		
220					225					230					235		
aat	gat	gcc	gat	ggt	atg	aaa	gag	ttg	tac	caa	ggt	aaa	tgt	gcg	ttg		1851
Asn	Asp	Ala	Asp	Gly	Met	Lys	Glu	Leu	Tyr	Gln	Gly	Lys	Cys	Ala	Leu		
				240					245					250			
acc	acc	acg	aac	cag	cag	atg	ttt	att	ctt	gcc	ggt	tcc	act	ttg	ggc		1899
Thr	Thr	Thr	Asn	Gln	Gln	Met	Phe	Ile	Leu	Ala	Gly	Ser	Thr	Leu	Gly		
			255					260					265				
ggt	ggt	acc	act	gtt	aac	tgg	tct	gct	tgt	ctt	aaa	aca	cca	ttt	aaa		1947
Gly	Gly	Thr	Thr	Val	Asn	Trp	Ser	Ala	Cys	Leu	Lys	Thr	Pro	Phe	Lys		
		270					275					280					
gtg	cgt	aag	gag	tgg	tac	gac	gag	ttt	ggt	ctt	gaa	ttt	gct	gcc	gac		1995
Val	Arg	Lys	Glu	Trp	Tyr	Asp	Glu	Phe	Gly	Leu	Glu	Phe	Ala	Ala	Asp		
	285					290					295						
gaa	gcc	tac	gac	aaa	gca	caa	gac	tat	gtt	tgg	aaa	caa	atg	ggc	gct		2043
Glu	Ala	Tyr	Asp	Lys	Ala	Gln	Asp	Tyr	Val	Trp	Lys	Gln	Met	Gly	Ala		
300					305					310				315			
tct	acc	gaa	gga	atc	act	cac	tct	ttg	gcg	aac	gcg	gtt	gtg	gtt	gaa		2091
Ser	Thr	Glu	Gly	Ile	Thr	His	Ser	Leu	Ala	Asn	Ala	Val	Val	Val	Glu		
				320					325					330			
gga	ggt	aag	aag	ttg	ggt	tac	aag	agc	aag	gaa	atc	gag	cag	aac	aat		2139
Gly	Gly	Lys	Lys	Leu	Gly	Tyr	Lys	Ser	Lys	Glu	Ile	Glu	Gln	Asn	Asn		
			335					340					345				
ggt	ggc	cat	cct	gac	cac	ccc	tgt	ggt	ttc	tgt	tac	ttg	ggc	tgt	aag		2187
Gly	Gly	His	Pro	Asp	His	Pro	Cys	Gly	Phe	Cys	Tyr	Leu	Gly	Cys	Lys		

ES 2 379 776 T3

350	355	360	
tac ggt att aag cag ggt tct gtg aat aac tgg ttt aga gac gca gct			2235
Tyr Gly Ile Lys Gln Gly Ser Val Asn Asn Trp Phe Arg Asp Ala Ala			
365	370	375	
gcc cac ggg tcc aag ttc atg caa caa gtc aga gtt gtg caa atc ctc			2283
Ala His Gly Ser Lys Phe Met Gln Gln Val Arg Val Val Gln Ile Leu			
380	385	390	395
cac aat aaa ggc gtc gct tat ggc atc ttg tgt gag gat gtc gag acc			2331
His Asn Lys Gly Val Ala Tyr Gly Ile Leu Cys Glu Asp Val Glu Thr			
	400	405	410
gga gtc aaa ttc act atc agt ggc ccc aaa aag ttt gtt gtt tct gca			2379
Gly Val Lys Phe Thr Ile Ser Gly Pro Lys Lys Phe Val Val Ser Ala			
	415	420	425
ggt tct ttg aac acg cca acg gtg ttg acc aac tcc gga ttc aag aac			2427
Gly Ser Leu Asn Thr Pro Thr Val Leu Thr Asn Ser Gly Phe Lys Asn			
	430	435	440
aaa cac atc ggt aag aac ttg acg ttg cac cca gtt tcg acc gtg ttt			2475
Lys His Ile Gly Lys Asn Leu Thr Leu His Pro Val Ser Thr Val Phe			
	445	450	455
ggt gac ttt ggc aga gac gtg caa gcc gac cat ttc cac aaa tct att			2523
Gly Asp Phe Gly Arg Asp Val Gln Ala Asp His Phe His Lys Ser Ile			
460	465	470	475
atg act tcg ctc tgt tac gaa gtc gct gac ttg gac ggc aag ggc cac			2571
Met Thr Ser Leu Cys Tyr Glu Val Ala Asp Leu Asp Gly Lys Gly His			
	480	485	490
gga tgc aga atc gag acc atc ttg aac gct cca ttc atc caa gct tct			2619
Gly Cys Arg Ile Glu Thr Ile Leu Asn Ala Pro Phe Ile Gln Ala Ser			
	495	500	505
ttg ttg cca tgg aga gga agc gac gag gtc aga aga gac ttg ttg cgt			2667
Leu Leu Pro Trp Arg Gly Ser Asp Glu Val Arg Arg Asp Leu Leu Arg			
	510	515	520
tac aac aac atg gtg gcc atg ttg ctt atc acc cgt gac acc acc agt			2715
Tyr Asn Asn Met Val Ala Met Leu Leu Ile Thr Arg Asp Thr Thr Ser			
	525	530	535
ggt tca gtt tct gct gac cca aag aag ccc gac gct ttg att gtc gac			2763
Gly Ser Val Ser Ala Asp Pro Lys Lys Pro Asp Ala Leu Ile Val Asp			
540	545	550	555
tat gac atc aac aag ttt gac aag aat gcc atc ttg caa gct ttc ttg			2811
Tyr Asp Ile Asn Lys Phe Asp Lys Asn Ala Ile Leu Gln Ala Phe Leu			
	560	565	570
atc acc tcc gac atg ttg tac atc gaa ggt gcc aag aga atc ctc agt			2859
Ile Thr Ser Asp Met Leu Tyr Ile Glu Gly Ala Lys Arg Ile Leu Ser			
	575	580	585

ES 2 379 776 T3

cca cag gca tgg gtg cca atc ttt gag tcg aac aag cca aag gag caa	2907
Pro Gln Ala Trp Val Pro Ile Phe Glu Ser Asn Lys Pro Lys Glu Gln	
590 595 600	
aga aca atc aag gac aag gac tat gtc gaa tgg aga gcc aag gct gcc	2955
Arg Thr Ile Lys Asp Lys Asp Tyr Val Glu Trp Arg Ala Lys Ala Ala	
605 610 615	
aag ata cct ttc gac acc tac ggt tct gcc tat ggg tcc gca cat caa	3003
Lys Ile Pro Phe Asp Thr Tyr Gly Ser Ala Tyr Gly Ser Ala His Gln	
620 625 630 635	
atg tcc acc tgt cgt atg tcc gga aag ggt cct aaa tac ggc gcc gtt	3051
Met Ser Thr Cys Arg Met Ser Gly Lys Gly Pro Lys Tyr Gly Ala Val	
640 645 650	
gat acc gat ggt aga ttg ttt gaa tgt tcg aat gtc tat gtt gct gat	3099
Asp Thr Asp Gly Arg Leu Phe Glu Cys Ser Asn Val Tyr Val Ala Asp	
655 660 665	
gct agt gtt ttg cct act gcc agc ggt gcc aac cca atg atc tcc acc	3147
Ala Ser Val Leu Pro Thr Ala Ser Gly Ala Asn Pro Met Ile Ser Thr	
670 675 680	
atg acg ttt gct aga cag att gcg tta ggt ttg gct gac tct ttg aag	3195
Met Thr Phe Ala Arg Gln Ile Ala Leu Gly Leu Ala Asp Ser Leu Lys	
685 690 695	
acc aaa ccc aag ttg tag agagagacag aaatacgaca cttatatact	3243
Thr Lys Pro Lys Leu	
700	
agatgtatct tacaatttat attttcgatg atggctttta ctatctocta tgttacacta	3303
taatgacatc accacatctt ctactactgt ctccagtatc ctccttgctg ttgaccgdat	3363
ccaccagcct gttggttgaa ccccgtgaac tgtggttgct gttgagcgta ccccacgtta	3423
gtgaactgcg gttggtgggt aaactgctgt acgggctggt gttgctggtg ctggtggtgc	3483
tggtggtgct gttggtgctg ttggtgctgt tgttggtggt gtcccgttggt ctggttgctac	3543
aacgacatga tgttctgctt gtttgtctgc tgggcaacca actgtggggtt attcatctgc	3603
atcaactgct gctggtgctg agggttggtt ggatccaagt actcctgccc gttggcgctg	3663
atataagaaa tctgccccgt gactgggtca gtgtactggt atatctgtgg catgccaccc	3723
gcttgtgcag gcatgccggt tgccaatggc	3753

<210> 12

<211> 704

<212> PRT

<213> *Candida tropicalis*

<400> 12

ES 2 379 776 T3

Met Asn Thr Phe Leu Pro Asp Val Leu Glu Tyr Lys His Val Asp Thr  
 1 5 10 15

Leu Leu Leu Leu Cys Asp Gly Ile Ile His Glu Thr Thr Val Asp Gln  
 20 25 30

Ile Arg Asp Ala Ile Ala Pro Asp Phe Pro Glu Asp Gln Tyr Glu Glu  
 35 40 45

Tyr Leu Lys Thr Phe Thr Lys Pro Ser Glu Thr Pro Gly Phe Arg Glu  
 50 55 60

Ala Val Tyr Asp Thr Ile Asn Ser Thr Pro Thr Glu Ala Val His Met  
 65 70 75 80

Cys Ile Val Leu Thr Thr Ala Leu Asp Ser Arg Ile Leu Ala Pro Thr  
 85 90 95

Leu Thr Asn Ser Leu Thr Pro Ile Lys Asp Met Thr Leu Lys Glu Arg  
 100 105 110

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

Arg Leu Phe Arg Leu Ile Ser Ser Leu Thr Leu Thr Thr Phe Thr Arg  
 130 135 140

Leu Ala Ser Asp Leu His Leu Arg Ala Ile His Tyr Pro Gly Arg Asp  
 145 150 155 160

Leu Arg Glu Lys Ala Tyr Glu Thr Gln Val Val Asp Pro Phe Arg Tyr  
 165 170 175

Ser Phe Met Glu Lys Pro Lys Phe Asp Gly Thr Glu Leu Tyr Leu Pro  
 180 185 190

Asp Ile Asp Val Ile Ile Ile Gly Ser Gly Ala Gly Ala Gly Val Met  
 195 200 205

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



ES 2 379 776 T3

Gly Lys Tyr Phe Ser Asn Ser Glu Leu Asn Phe Asn Asp Ala Asp Gly  
 225 230 235 240

Met Lys Glu Leu Tyr Gln Gly Lys Cys Ala Leu Thr Thr Thr Asn Gln  
 245 250 255

Gln Met Phe Ile Leu Ala Gly Ser Thr Leu Gly Gly Gly Thr Thr Val  
 260 265 270

Asn Trp Ser Ala Cys Leu Lys Thr Pro Phe Lys Val Arg Lys Glu Trp  
 275 280 285

Tyr Asp Glu Phe Gly Leu Glu Phe Ala Ala Asp Glu Ala Tyr Asp Lys  
 290 295 300

Ala Gln Asp Tyr Val Trp Lys Gln Met Gly Ala Ser Thr Glu Gly Ile  
 305 310 315 320

Thr His Ser Leu Ala Asn Ala Val Val Val Glu Gly Gly Lys Lys Leu  
 325 330 335

Gly Tyr Lys Ser Lys Glu Ile Glu Gln Asn Asn Gly Gly His Pro Asp  
 340 345 350

His Pro Cys Gly Phe Cys Tyr Leu Gly Cys Lys Tyr Gly Ile Lys Gln  
 355 360 365

Gly Ser Val Asn Asn Trp Phe Arg Asp Ala Ala Ala His Gly Ser Lys  
 370 375 380

Phe Met Gln Gln Val Arg Val Val Gln Ile Leu His Asn Lys Gly Val  
 385 390 395 400

Ala Tyr Gly Ile Leu Cys Glu Asp Val Glu Thr Gly Val Lys Phe Thr  
 405 410 415

Ile Ser Gly Pro Lys Lys Phe Val Val Ser Ala Gly Ser Leu Asn Thr  
 420 425 430

Pro Thr Val Leu Thr Asn Ser Gly Phe Lys Asn Lys His Ile Gly Lys  
 435 440 445

ES 2 379 776 T3

Asn Leu Thr Leu His Pro Val Ser Thr Val Phe Gly Asp Phe Gly Arg  
 450 455 460

Asp Val Gln Ala Asp His Phe His Lys Ser Ile Met Thr Ser Leu Cys  
 465 470 475 480

Tyr Glu Val Ala Asp Leu Asp Gly Lys Gly His Gly Cys Arg Ile Glu  
 485 490 495

Thr Ile Leu Asn Ala Pro Phe Ile Gln Ala Ser Leu Leu Pro Trp Arg  
 500 505 510

Gly Ser Asp Glu Val Arg Arg Asp Leu Leu Arg Tyr Asn Asn Met Val  
 515 520 525

Ala Met Leu Leu Ile Thr Arg Asp Thr Thr Ser Gly Ser Val Ser Ala  
 530 535 540

Asp Pro Lys Lys Pro Asp Ala Leu Ile Val Asp Tyr Asp Ile Asn Lys  
 545 550 555 560

Phe Asp Lys Asn Ala Ile Leu Gln Ala Phe Leu Ile Thr Ser Asp Met  
 565 570 575

Leu Tyr Ile Glu Gly Ala Lys Arg Ile Leu Ser Pro Gln Ala Trp Val  
 580 585 590

Pro Ile Phe Glu Ser Asn Lys Pro Lys Glu Gln Arg Thr Ile Lys Asp  
 595 600 605

Lys Asp Tyr Val Glu Trp Arg Ala Lys Ala Ala Lys Ile Pro Phe Asp  
 610 615 620

Thr Tyr Gly Ser Ala Tyr Gly Ser Ala His Gln Met Ser Thr Cys Arg  
 625 630 635 640

Met Ser Gly Lys Gly Pro Lys Tyr Gly Ala Val Asp Thr Asp Gly Arg  
 645 650 655

Leu Phe Glu Cys Ser Asn Val Tyr Val Ala Asp Ala Ser Val Leu Pro  
 660 665 670

Thr Ala Ser Gly Ala Asn Pro Met Ile Ser Thr Met Thr Phe Ala Arg

675

680

685

Gln Ile Ala Leu Gly Leu Ala Asp Ser Leu Lys Thr Lys Pro Lys Leu  
 690 695 700

<210> 13

<211> 8

<212> PRT

<213> *Candida tropicalis*

<400> 13

Cys Gly Phe Cys Tyr Leu Gly Cys  
 1 5

<210> 14

<211> 12

<212> PRT

<213> *Candida tropicalis*

<400> 14

Ile Ile Gly Ser Gly Ala Gly Ala Gly Val Met Ala  
 1 5 10

<210> 15

<211> 12

<212> PRT

<213> *Candida tropicalis*

<220>

<221> característica nueva

<222> (6)..(6)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 15

Ile Ile Gly Ser Gly Xaa Gly Ala Gly Val Val Ala  
1 5 10

<210> 16

<211> 12

<212> PRT

<213> *Candida tropicalis*

<400> 16

Ile Ile Gly Ser Gly Ala Gly Ala Gly Val Val Ala  
1 5 10

<210> 17

<211> 12

<212> PRT

<213> *Candida albicans*

<400> 17

Ile Ile Gly Ser Gly Ala Gly Ser Gly Val Val Ala  
1 5 10

<210> 18

<211> 8

<212> PRT

<213> *Candida tropicalis*

<400> 18

Ala Gly Ser Thr Phe Gly Gly Gly  
1 5

<210> 19

<211> 8

<212> PRT

<213> *Candida tropicalis*

<400> 19

Ala Gly Ser Thr Leu Gly Gly Gly  
1 5

<210> 20

<211> 9

<212> PRT

<213> *Candida tropicalis*

<400> 20

Asn Trp Ser Ala Cys Leu Lys Thr Pro  
1 5

<210> 21

<211> 9

<212> PRT

<213> *Candida albicans*

<400> 21

Asn Trp Ser Ala Cys Ile Lys Thr Pro  
1 5

<210> 22

<211> 8

<212> PRT

<213> *Candida tropicalis*

<220>

<221> característica nueva

<222> (3)..(3)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 22

Cys Gly Xaa Cys His Leu Gly Cys  
1 5

<210> 23

<211> 8

<212> PRT

<213> *Candida tropicalis*

<400> 23

Cys Gly Phe Cys His Leu Gly Cys  
1 5

<210> 24

<211> 11

<212> PRT

<213> *Candida tropicalis*

<220>

<221> característica nueva

<222> (3)..(3)

<223> Xaa puede ser cualquier aminoácido de origen natural

<220>

<221> característica nueva

<222> (6)..(6)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 24

Ile Gly Xaa Asn Leu Xaa Leu His Pro Val Ser  
1 5 10

<210> 25

<211> 11

<212> PRT

<213> *Candida tropicalis*

<400> 25

Ile Gly Lys Asn Leu Thr Leu His Pro Val Ser  
1 5 10

<210> 26

<211> 11

<212> PRT

<213> *Candida albicans*

<400> 26

Ile Gly Ala Asn Leu Thr Leu His Pro Val Thr  
1 5 10

<210> 27

<211> 12

<212> PRT

<213> *Candida tropicalis*

<220>

<221> característica nueva

<222> (7)..(7)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 27

Ser Ala His Gln Met Ser Xaa Cys Arg Met Ser Gly  
1 5 10

<210> 28

<211> 12

<212> PRT

<213> *Candida tropicalis*

<400> 28

Ser Ala His Gln Met Ser Thr Cys Arg Met Ser Gly  
1 5 10

<210> 29

<211> 12

<212> PRT

<213> *Candida tropicalis*

<400> 29

Ser Ala His Gln Met Ser Ser Cys Arg Met Ser Gly  
1 5 10

<210> 30

<211> 9

<212> PRT

<213> *Candida tropicalis*

<220>



<221> característica nueva

<222> (6)..(6)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 30

Pro Thr Ala Ser Gly Xaa Asn Pro Met  
1 . 5

<210> 31

<211> 9

<212> PRT

<213> *Candida tropicalis*

<400> 31

Pro Thr Ala Ser Gly Ala Asn Pro Met  
1 - 5

<210> 32

<211> 24

<212> ADN

<213> *Candida tropicalis*

<220>

<221> característica nueva

<222> (3)..(3)

<223> y e s c ó t

<220>

<221> característica nueva

<222> (6)..(6)

<223> n e s a, c, g, ó t

<220>

<221> característica nueva

<222> (9)..(9)

<223> y es c ó t

<220>

<221> característica nueva

<222> (12)..(12)

<223> y es c ó t

<220>

<221> característica nueva

<222> (15)..(15)

<223> y es c ó t

<220>

<221> característica nueva

<222> (16)..(16)

<223> y es c ó t

<220>

<221> característica nueva

<222> (18)..(18)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (21)..(21)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (24)..(24)

<223> y es c ó t

<400> 32

tgyggnttyt gytayytnng ntgy 24

<210> 33

<211> 36

<212> ADN

<213> *Candida tropicalis*

<220>

<221> característica nueva

<222> (3)..(3)

<223> h e s a ó t ó c

<220>

<221> característica nueva

<222> (6)..(6)

<223> h e s a ó t ó c

<220>

<221> característica nueva

<222> (9)..(9)

<223> n e s a, c, g, ó t

<220>

<221> característica nueva

<222> (10)..(10)

<223> w e s a ó t

<220>

<221> característica nueva

<222> (11)..(11)

<223> s e s g ó c

<220>

<221> característica nueva

<222> (12)..(12)

<223> n e s a, c, g, ó t

<220>

<221> característica nueva

<222> (15)..(15)

<223> n e s a, c, g, ó t

<220>

<221> característica nueva

<222> (18)..(18)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (21)..(21)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (24)..(24)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (27)..(27)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (30)..(30)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (36)..(36)

<223> n es a, c, g, ó t

<400> 33

athathggnw sngngcngg ngcngngtn atggcn 36

<210> 34

<211> 24

<212> ADN

<213> *Candida tropicalis*

<220>

<221> característica nueva

<222> (3)..(3)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (6)..(6)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (7)..(7)

<223> w es a ó t

<220>

<221> característica nueva

<222> (8)..(8)

<223> s es g ó c

<220>

<221> característica nueva

<222> (9)..(9)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (12)..(12)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (13)..(13)

<223> y es c ó t

<220>

<221> característica nueva

<222> (15)..(15)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (18)..(18)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (21)..(21)

<223> n es a, c, g, ó t

<220>

<221> característica nueva

<222> (24)..(24)

<223> n es a, c, g, ó t

<400> 34

gcnggnwsna cnytngngg nggn 24

<210> 35

<211> 18

<212> ADN

<213> *Candida tropicalis*

<400> 35

tcgtggcgtg actctcct 18

<210> 36

<211> 18

<212> ADN

<213> *Candida cloacae*

<400> 36

tcatggagag actctcct 18

<210> 37

<211> 18

<212> ADN  
<213> *Candida cloacae*  
<400> 37  
tcatggagag actctcca 18  
<210> 38  
<211> 16  
<212> ADN  
<213> *Candida tropicalis*  
<400> 38  
ctggtgctgg tgtagt 16  
<210> 39  
<211> 16  
<212> ADN  
<213> *Candida cloacae*  
<400> 39  
caggagcagg tgtggt 16  
<210> 40  
<211> 16  
<212> ADN  
<213> *Candida cloacae*  
<400> 40  
cgggagcagg agtggt 16  
<210> 41  
<211> 19  
<212> ADN  
<213> *Candida tropicalis*  
<400> 41  
ttgtaccca tgctgtgg 19  
<210> 42  
<211> 19

<212> ADN

<213> *Candida cloacae*

<400> 42

ttgtaccca agcttgtag 19

<210> 43

<211> 19

<212> ADN

<213> *Candida cloacae*

<400> 43

ttgtaccca agcttgtag 19

<210> 44

<211> 18

<212> ADN

<213> *Candida tropicalis*

<400> 44

tcgtggcgtg actcccct 18

<210> 45

<211> 18

<212> ADN

<213> *Candida tropicalis*

<400> 45

tcttggcgtg attccccg 18

<210> 46

<211> 18

<212> ADN

<213> *Candida tropicalis*

<400> 46

gcctggcgtg attccccg 18

<210> 47

<211> 16



<212> ADN

<213> *Candida tropicalis*

<400> 47

ccggtgctgg tgcgt 16

<210> 48

<211> 16

<212> ADN

<213> *Candida tropicalis*

<400> 48

ccggtgctgg tgtcat 16

<210> 49

<211> 16

<212> ADN

<213> *Candida tropicalis*

<400> 49

ccggtgctgg tgtcat 16

<210> 50

<211> 19

<212> ADN

<213> *Candida tropicalis*

<400> 50

ttggcaccca tgggtggg 19

<210> 51

<211> 19

<212> ADN

<213> *Candida tropicalis*

<400> 51

ttggcaccca tggctgtg 19

<210> 52

<211> 19

<212> ADN

<213> *Candida tropicalis*

<400> 52

ttggcaccca tgcctgtgg 19

<210> 53

<211> 26

<212> ADN

<213> *Candida tropicalis*

<400> 53

ccgaattcga catggctcca ttttg 26

<210> 54

<211> 30

<212> ADN

<213> *Candida tropicalis*

<400> 54

ccggatccat tactacaact tggccttgg 30

<210> 55

<211> 26

<212> ADN

<213> *Candida tropicalis*

<400> 55

ccagtgaatt cagatgaata ccttct 26

<210> 56

<211> 26

<212> ADN

<213> *Candida tropicalis*

<400> 56

ccggatcccc gtctcactac aacttg 26

<210> 57

<211> 18

<212> ADN

<213> *Candida tropicalis*

<400> 57

atcaacgcca cccaacc 18

<210> 58

<211> 36

<212> ADN

<213> *Candida tropicalis*

<400> 58

ggtttctcca taaacgagta cctgaaagg tcaacc 36

<210> 59

<211> 18

<212> ADN

<213> *Candida tropicalis*

<400> 59

atctgtctag caaaggtc 18

<210> 60

<211> 36

<212> ADN

<213> *Candida tropicalis*

<400> 60

ggttgaccct tcaggtact cgttatgga gaaacc 36

<210> 61

<211> 31

<212> ADN

<213> *Candida tropicalis*

<400> 61

ccttaattaa tgcatactcg gagcatatcg c 31

<210> 62

<211> 28

<212> ADN  
 <213> *Candida tropicalis*  
 <400> 62  
 ccttaattaa tgggcggaat caagtgcc 28  
 <210> 63  
 <211> 30  
 <212> ADN  
 <213> *Candida tropicalis*  
 <400> 63  
 ccttaattaa ttcaccaag tacgagaacg 30  
 <210> 64  
 <211> 28  
 <212> ADN  
 <213> *Candida tropicalis*  
 <400> 64  
 ccttaattaa gacgcaagca caggtgcc 28  
 <210> 65  
 <211> 29  
 <212> ADN  
 <213> *Candida tropicalis*  
 <400> 65  
 ccttaattaa agtctccaag ttgaccgac 29  
 <210> 66  
 <211> 26  
 <212> ADN  
 <213> *Candida tropicalis*  
 <400> 66  
 aaatggagcc atggctgtga tgtgtg 26  
 <210> 67  
 <211> 29

<212> ADN  
 <213> *Candida tropicalis*  
 <400> 67  
 cacatcacga ccatggctcc attttgcc 29  
 <210> 68  
 <211> 28  
 <212> ADN  
 <213> *Candida tropicalis*  
 <400> 68  
 ccttaattaa tgggcggaat caagtgcc 28  
 <210> 69  
 <211> 29  
 <212> ADN  
 <213> *Candida tropicalis*  
 <400> 69  
 ccttaattaa agtctccaag ttgaccgac 29  
 <210> 70  
 <211> 26  
 <212> ADN  
 <213> *Candida tropicalis*  
 <400> 70  
 gaaggatttc atggtcgtga tgtgtg 26  
 <210> 71  
 <211> 29  
 <212> ADN  
 <213> *Candida tropicalis*  
 <400> 71  
 cacatcacga ccatgaatac cttcttgcc 29  
 <210> 72  
 <211> 28

<212> ADN

<213> *Candida tropicalis*

<400> 72

ccttaattaa gacgcaagca caggtgcc 28

#### Referencias

1. Kemp G. D., Dickinson M, Ratledge C (1988) "Inducible long chain alcohol oxidase from alkane-grown *Candida tropicalis*." *Appl. Microbiol. Biotechnol.* 29: 370 - 374.
2. Kemp G. D., Dickinson F. M., Ratledge C (1991) "Activity and substrate specificity of the fatty alcohol oxidase of *Candida tropicalis* in organic solvents." *Appl. Microbiol. Biotechnol.* 34: 441 - 445.
3. Dickinson F. M., Wadforth C (1992) "Purification and some properties of alcohol oxidase from alkane-grown *Candida tropicalis*." *Biochem. J.* 282: 325 - 331.
4. Vanhanen S, West M, Kroon JTM, Lindner N, Casey J, Cheng Q, Elborough K. M., Slabas A. R. (2000) "A consensus sequence for long-chain fatty-acid alcohol oxidases from *Candida* identifies a family of genes involved in lipidoxidation in yeast with homologues in plants and bacteria." *J. Biol. Chem.* 275: 4445 - 4452.
5. Blasig R, Mauersberger S, Riege R, Schunck W. H., Jockisch W, Franke P (1988) "Degradation of long-chain nalkanes by the yeast *Candida maltosa*. II. Oxidation of n-alkanes and intermediates using microsomal membrane fractions." *Appl. Microbiol. Biotechnol.* 28: 589 - 597.
6. Mauersberger S, Drechsler H, Oehme G, Müller HG (1992) "Substrate specificity and stereoselectivity of fatty alcohol oxidase from the yeast *Candida maltosa*." *Appl. Microbiol. Biotechnol.* 37: 66 - 73.
7. Ilchenko A. P., Tsfasman I. M. (1988) "Isolation and characterization of alcohol oxidase from higher alcohols of the yeast *Torulopsis candida* grown on hexadecane." *Biokhimiya* 53: 263 - 271.
8. Hommel R, Ratledge C (1990) "Evidence for two oxidases de alcohol graso in the biosurfactant-producing yeast *Candida (Torulopsis) bombicola*." *FEMS Microbiol. Lett.* 70: 183 - 186.
9. Hommel R, Lassner D, Weiss J, Kleber H. P. (1994) "The inducible microsomal fatty alcohol oxidase of *Candida (Torulopsis) apicola*." *Appl. Microbiol. Biotechnol.* 40: 729 - 734.
10. Kemp G. D., Dickinson F. M., Ratledge C (1990) "Light sensitivity of the n-alkane-induced fatty alcohol oxidase from *Candida tropicalis* and *Yarrowia lipolytica* aproximadamente" *Appl. Microbiol. Biotechnol.* 32: 461 - 464.
11. Eirich L. D. (1989) "Partial characterization of 12-hydroxylauric acid oxidase in *Candida tropicalis*" Topical Report Ei2/89.
12. Slabas A. R., Elborough K, Vanhanen S, West M, Cheng Q, Lindner N, Casey J, Sanglard D (1999) "Improvements in o relating to ácido graso metabolism" International Patent Application WO 99/47685.
13. Fürste J. P., Pansegrau W, Frank R, Blöcker H, Scholz P, Bagdasarian M, Lanka E. (1986) "Molecular cloning of the plasmid RP4 primase region in a multi-host- range tacP expression vector." *Gene* 48: 119 - 131.
14. Ueda T, Suzuki T, Yokogawa T, Nishikawa K, Watanabe K (1994) "Unique structure of new serine tRNAs responsible for decoding leucine codon CUG in various *Candida* species and their putative ancestral tRNA genes." *Biochimie* 76: 1217 - 1222.

**REIVINDICACIONES**

1. Una oxidasa de alcohol graso o un fragmento enzimáticamente activo de la misma que tiene una identidad de secuencia de aminoácidos superior al 90% cuando se la compara con la secuencia de aminoácidos expuesta en la SEQ ID NO: 2.
2. Una oxidasa de alcohol graso de acuerdo con la reivindicación 1, o un fragmento enzimáticamente activo de la misma que tiene una identidad de secuencia de aminoácidos superior al 95% cuando se la compara con la secuencia de aminoácidos expuesta en la SEQ ID NO: 2.
3. La oxidasa de alcohol graso de la reivindicación 1 ó 2 que comprende además un péptido firma que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 13.
4. Una molécula aislada de ácido nucleico que codifica una oxidasa de alcohol graso que tiene una secuencia de aminoácidos como la expuesta en la SEQ ID NO: 2, o un análogo, derivado o fragmento enzimáticamente activo de la misma que tiene una identidad de secuencia de aminoácidos superior al 90% cuando se la compara con la secuencia de aminoácidos expuesta en la SEQ ID NO: 2.
5. La molécula aislada de ácido nucleico de la reivindicación 4 que contiene una secuencia de nucleótidos como la expuesta en la SEQ ID NO: 1.
6. Un vector que comprende una de las moléculas aisladas de ácido nucleico seleccionadas de entre el grupo que consiste de moléculas aisladas de ácido nucleico de la reivindicación 4 o la reivindicación 5.
7. Los vectores de la reivindicación 6 en donde el vector es un vector plásmido, fagémido, fago, cósmido, o vector de ADN lineal.
8. Una célula huésped que comprende uno de los vectores de la reivindicación 7.
9. La célula huésped de la reivindicación 8 en donde la célula es una célula bacteriana, una célula de hongo, una célula de insecto, una célula de animal o una célula de una planta.
10. La célula de hongo de la reivindicación 9 en donde la célula de hongo es una célula de levadura seleccionada de entre el grupo que consiste de *Yarrowia sp.*, *Bebaromyces sp.*, *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.* y *Candida sp.*
11. Un método para la producción de una proteína FAO1 que tiene una secuencia de aminoácidos como la expuesta en la SEQ ID NO: 2, dicho método comprendiendo: la transformación de una célula huésped adecuada con una secuencia de ADN que codifica una proteína que tiene una secuencia de aminoácidos como la expuesta en la SEQ ID NO: 2, y el cultivo de la célula bajo condiciones que favorezcan la expresión de la proteína FAO1.

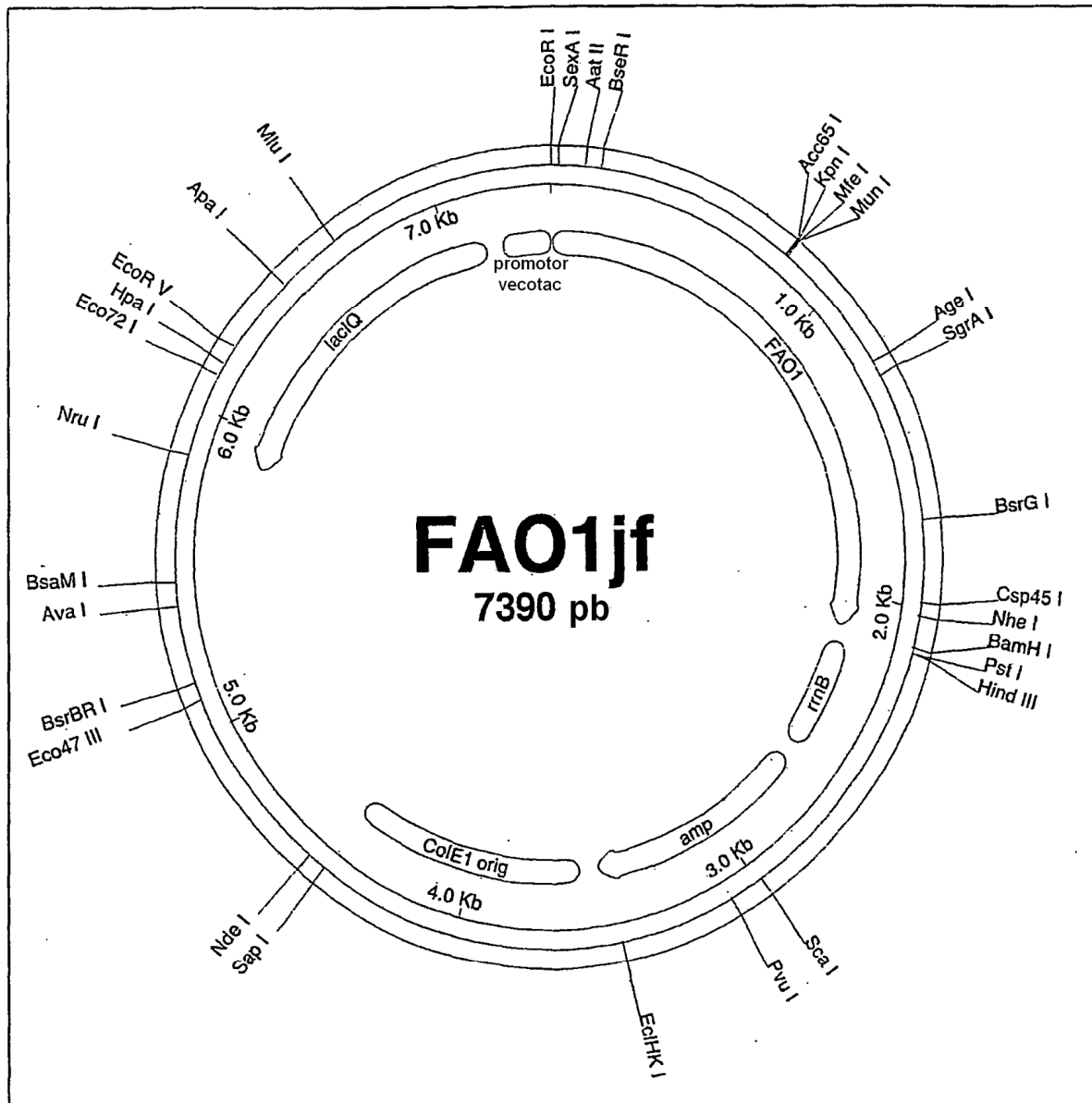


FIGURA 1



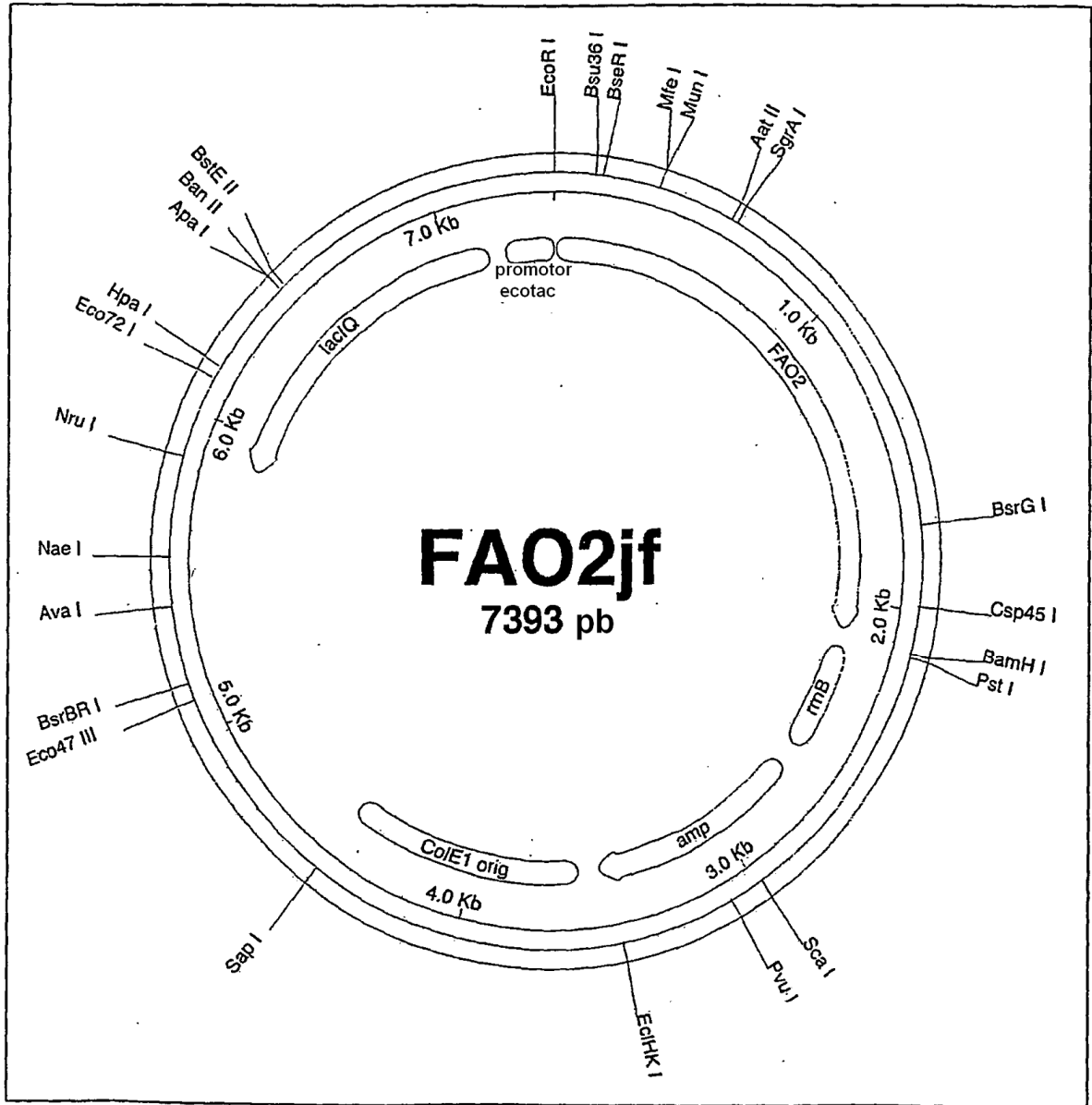
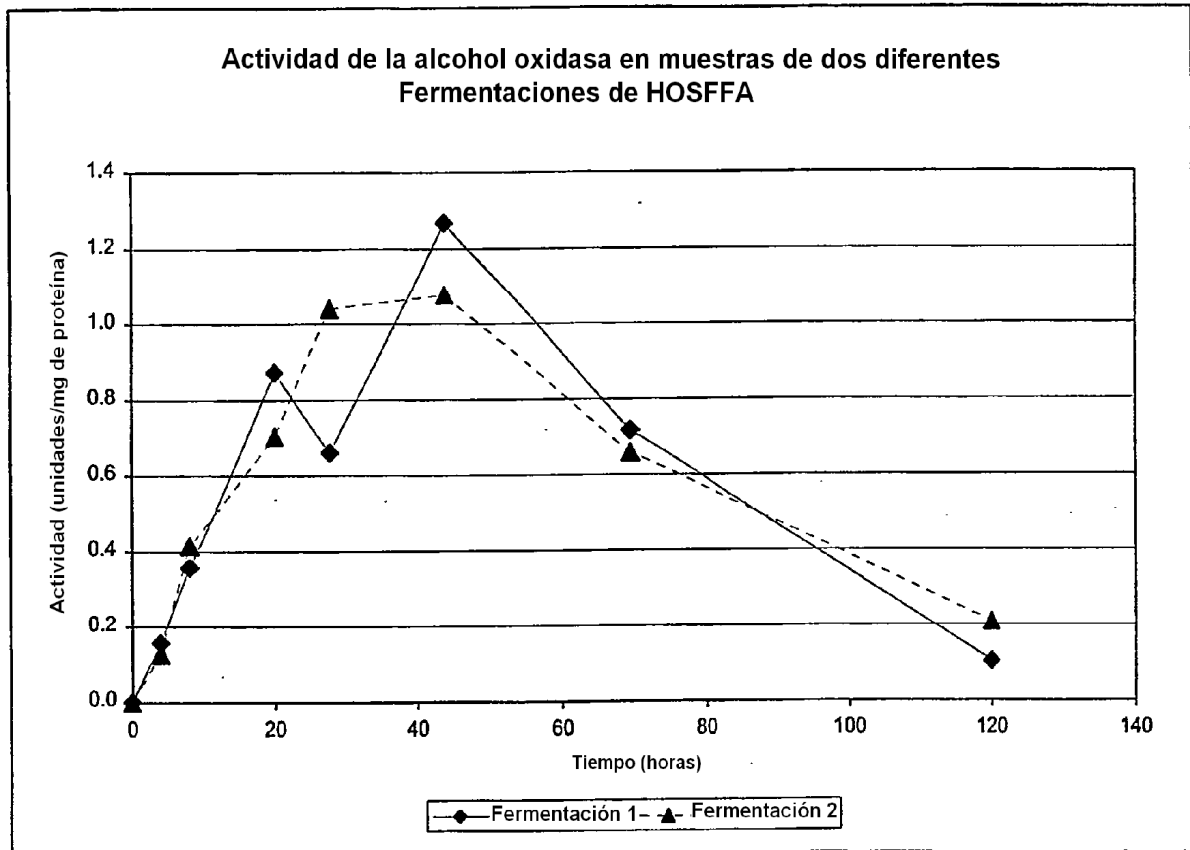


FIGURA 2



**FIGURA 3**

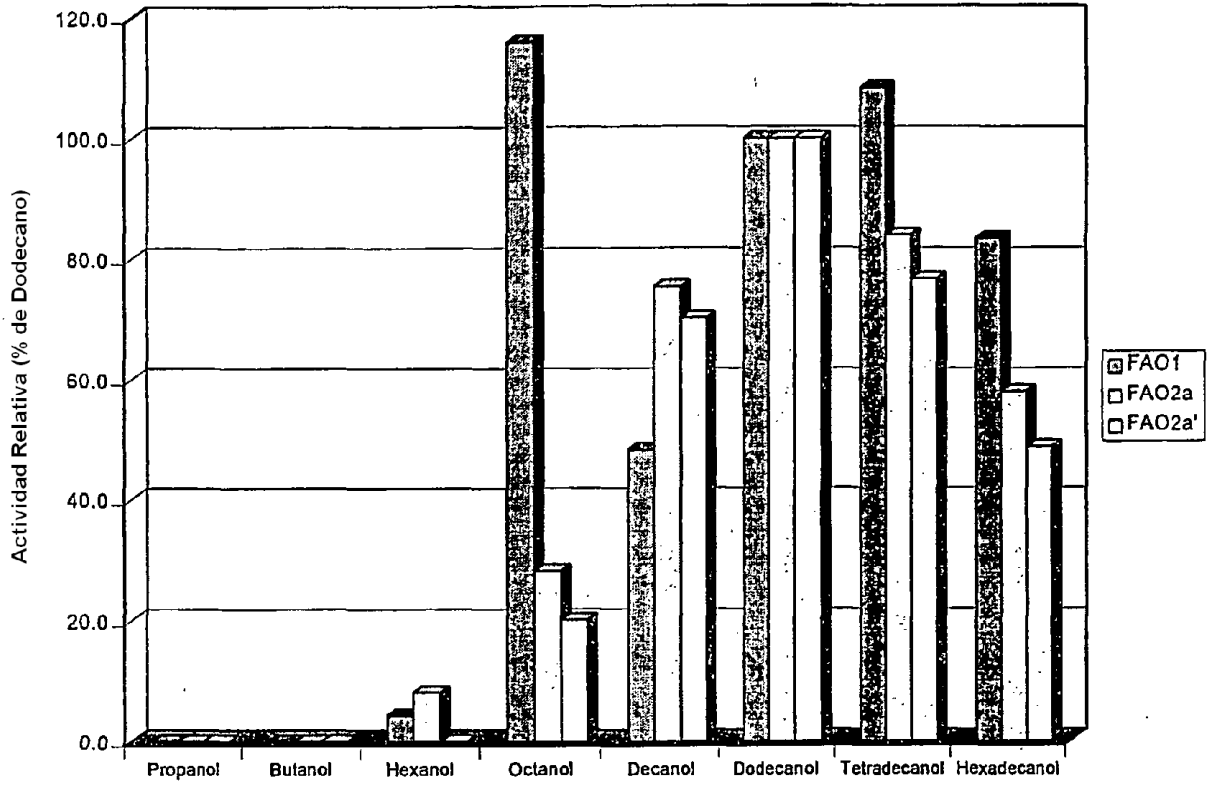


FIGURA 4

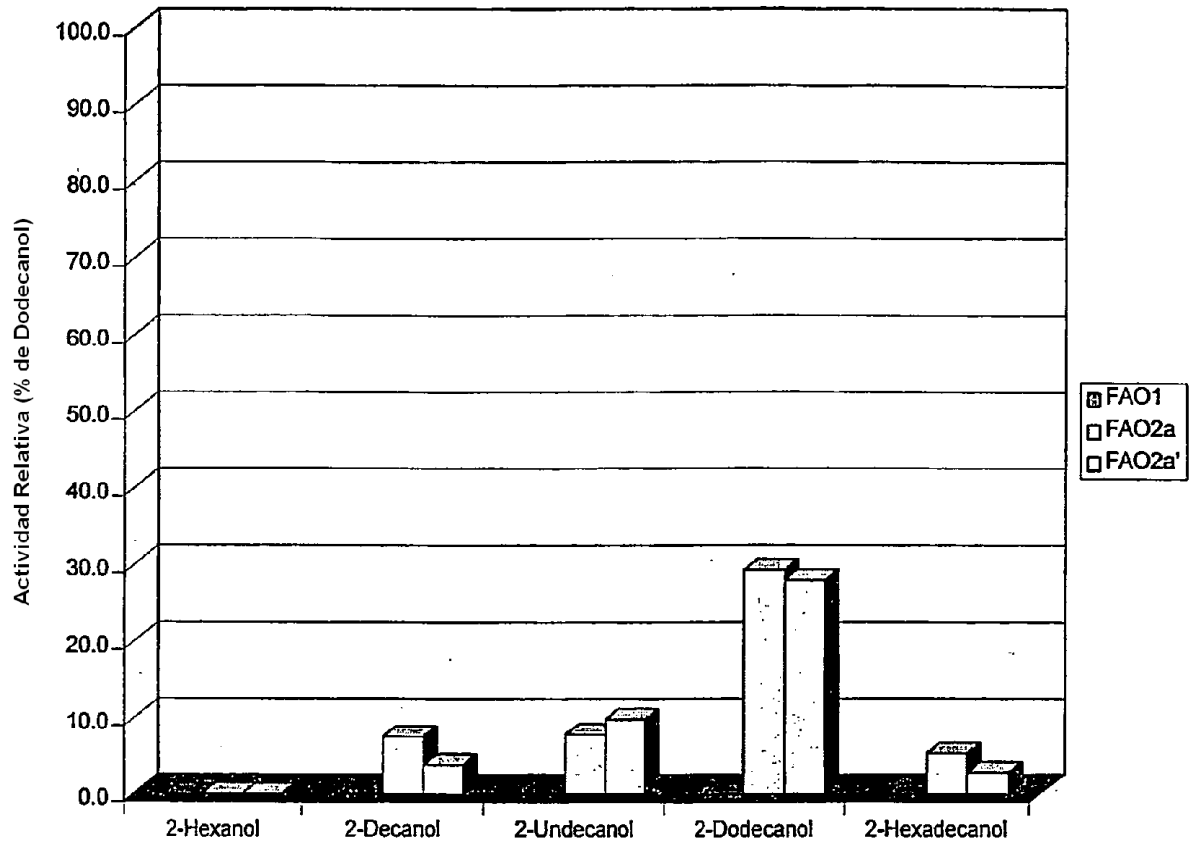


FIGURA 5

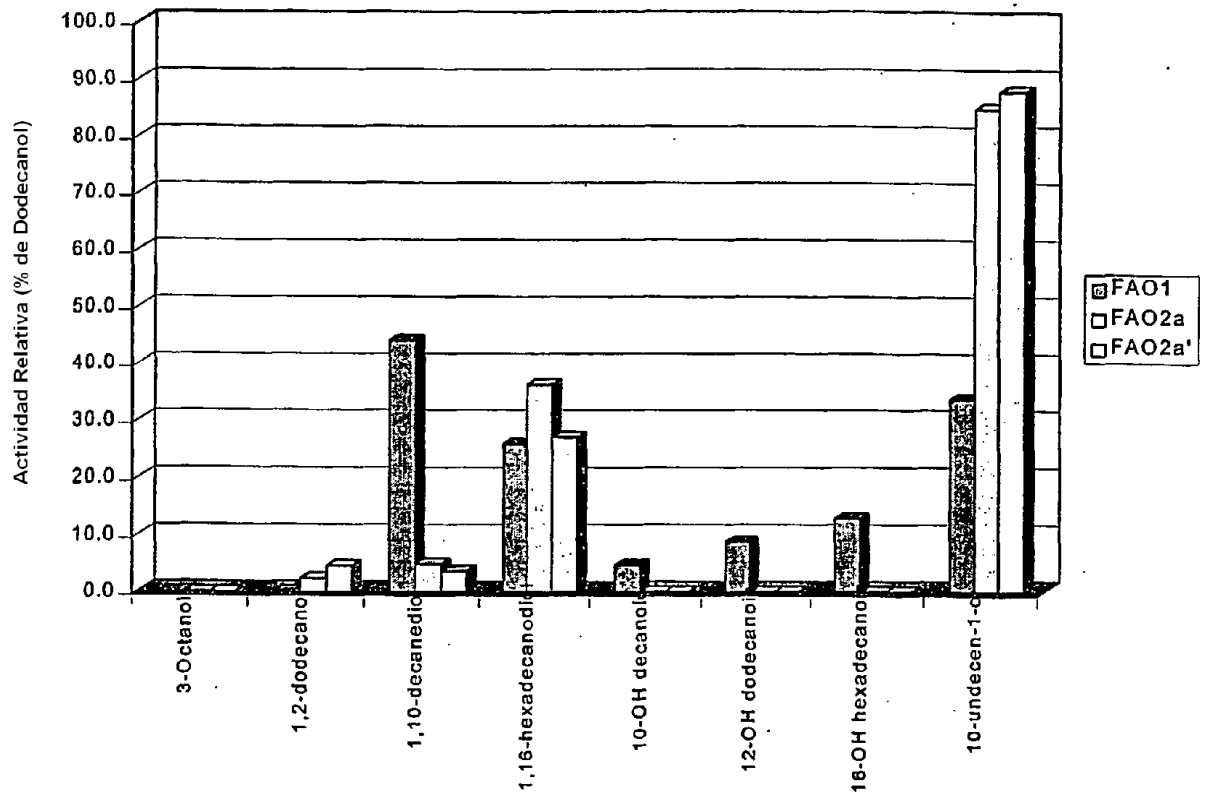


FIGURA 6

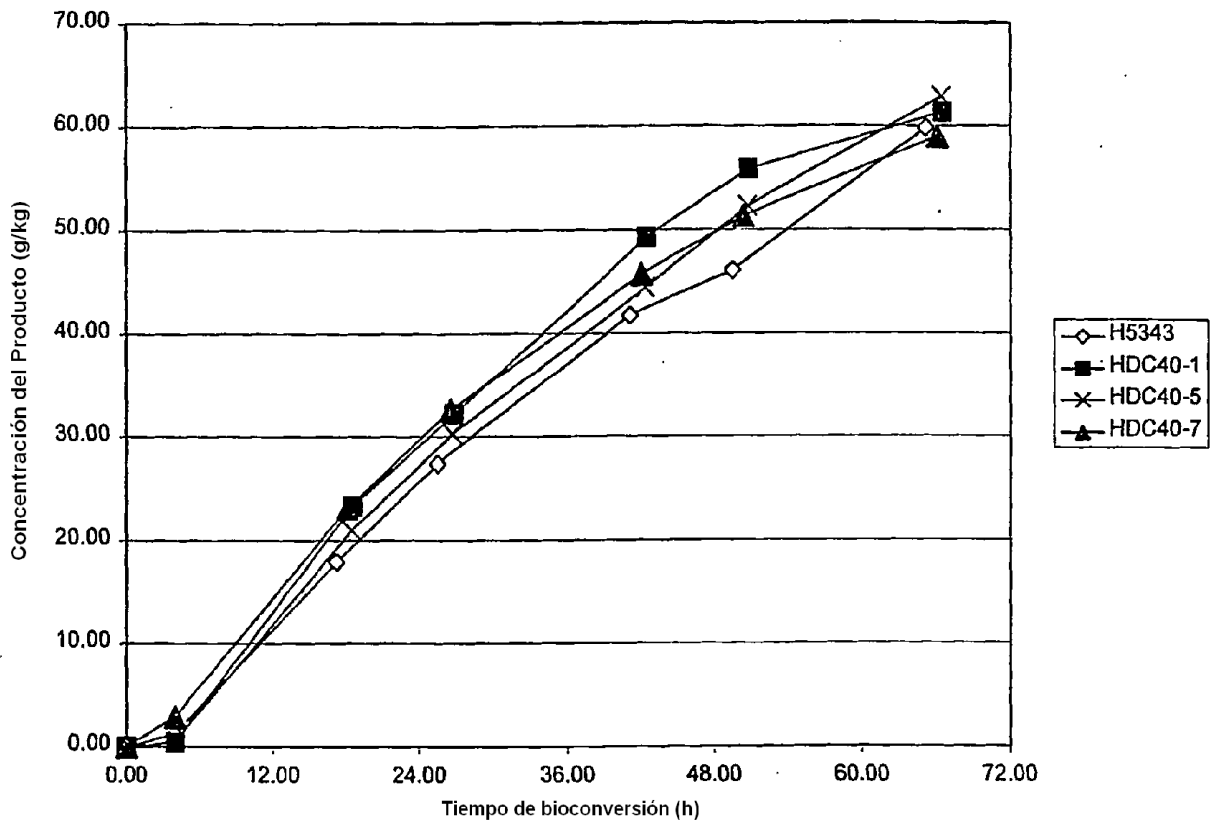


FIGURA 7

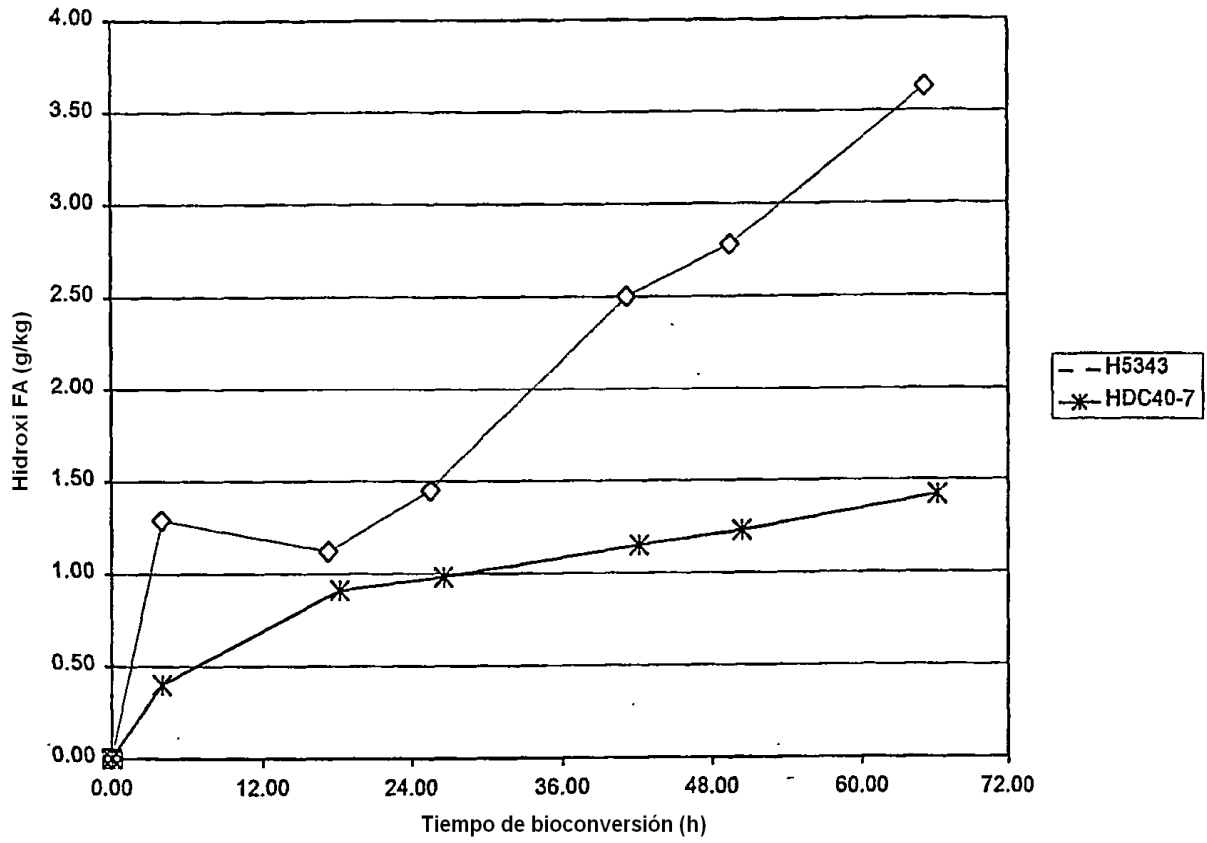


FIGURA 8

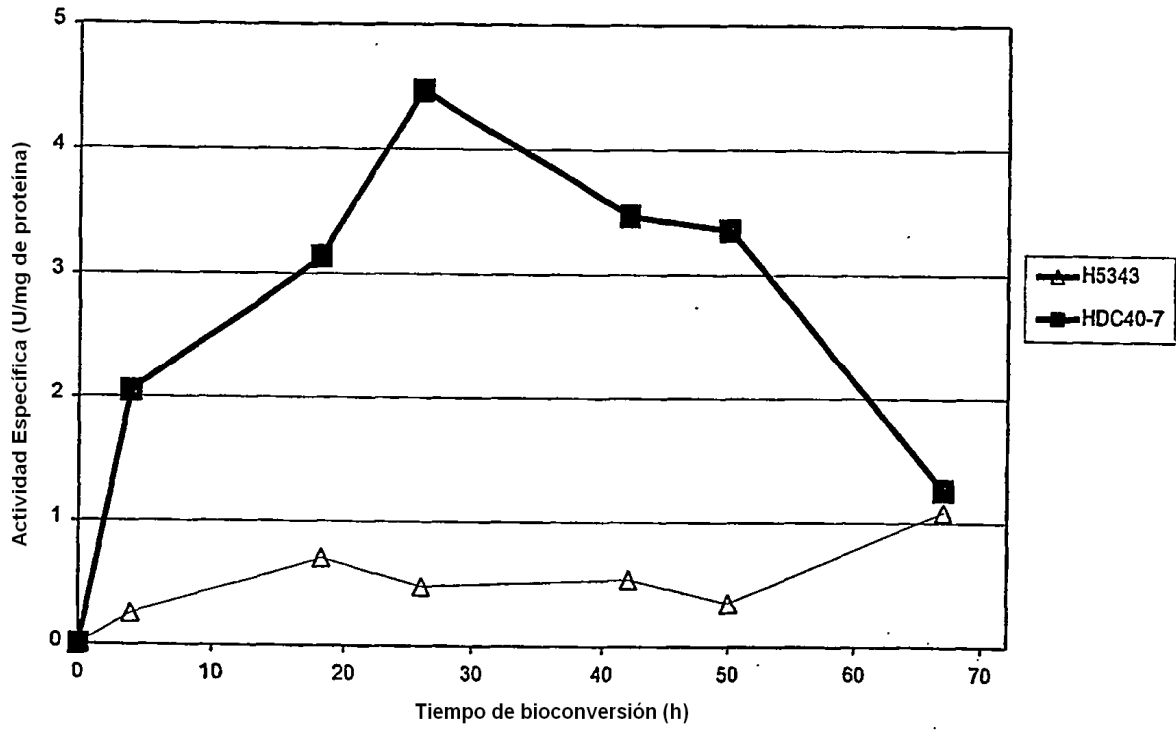


FIGURA 9



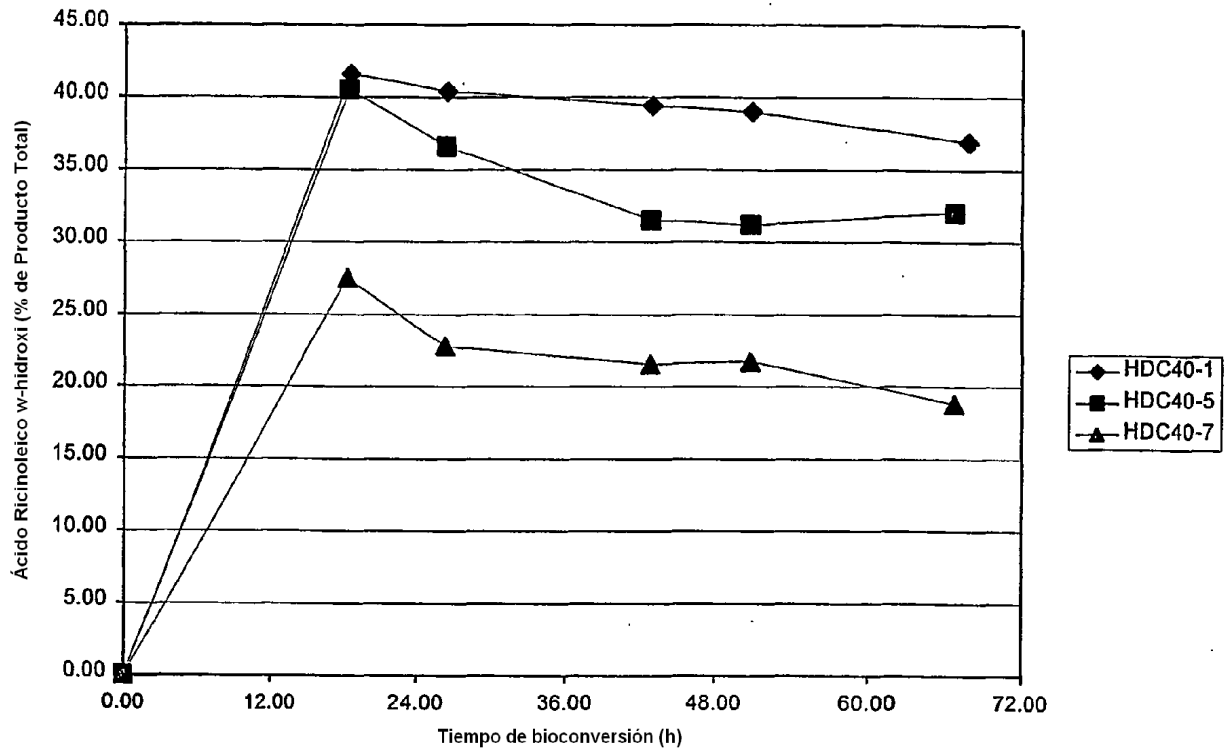


FIGURA 10