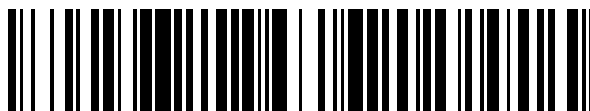


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 392 742**

51 Int. Cl.:
C07K 14/35 (2006.01)
C12N 9/02 (2006.01)
C12N 9/04 (2006.01)
C12N 15/63 (2006.01)
C12N 1/20 (2006.01)
C12Q 1/68 (2006.01)
C12N 15/11 (2006.01)
C02F 3/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06778881 .0**
96 Fecha de presentación: **18.07.2006**
97 Número de publicación de la solicitud: **1913021**
97 Fecha de publicación de la solicitud: **23.04.2008**

54 Título: **Polipéptidos que tienen una actividad en la vía de degradación del metil-terc.butil-éter (MTBE) y sus utilizaciones**

30 Prioridad:
18.07.2005 FR 0507577

45 Fecha de publicación de la mención BOPI:
13.12.2012

45 Fecha de la publicación del folleto de la patente:
13.12.2012

73 Titular/es:
**INSTITUT FRANÇAIS DU PÉTROLE (50.0%)
1 & 4 AVENUE DE BOIS-PRÉAU
92852 RUEIL MALMAISON CEDEX, FR y
CONSEIL NATIONAL DE RECHERCHES DU
CANADA (50.0%)**

72 Inventor/es:
**LOPES FERREIRA, NICOLAS;
LABBE, DIANE;
MACIEL, HÉLÉNA;
FAYOLLE-GUICHARD, FRANÇOISE;
MONOT, FRÉDÉRIC y
GREER, CHARLES W.**

74 Agente/Representante:
DE ELZABURU MÁRQUEZ, Alberto

ES 2 392 742 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

Polipéptidos que tienen una actividad en la vía de degradación del metil-*terc*.butil-éter (MTBE) y sus utilidades.

La presente solicitud se refiere al dominio de la microbiología, y más particularmente a la utilización de microorganismos para tratar efluentes contaminados por contaminantes químicos, principalmente el metil-*terc*-butil-éter (MTBE) y/o intermedios catabólicos de este compuesto.

La presente invención se refiere a nuevos polipéptidos y sus fragmentos, así como a los ácidos nucleicos que codifican dichos polipéptidos aislados de microorganismos capaces de metabolizar el MTBE, y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente el alcohol *terc*-butílico (TBA), el 2-metil-1,2-propanodiol (2-M-1,2-PD), el hidroxibutiraldehído, el ácido hidroxibutírico (HIBA). La invención se refiere igualmente a vectores de clonación y/o de expresión que comprenden dichos ácidos nucleicos, células bacterianas transformadas por dichos ácidos nucleicos o dichos vectores, así como sus utilidades.

La invención se refiere igualmente a un procedimiento de identificación de microorganismos capaces de metabolizar el MTBE, y/o al menos uno de los intermedios catabólicos del MTBE. Más particularmente, la presente invención se refiere a una nueva cepa bacteriana, *Mycobacterium austroafricanum*, registrada en la colección CNCM bajo el número I-3401.

La invención se refiere igualmente a un procedimiento de tratamiento de efluentes contaminados por el MTBE y/o al menos uno de los intermedios catabólicos del MTBE.

El compuesto químico MTBE se utiliza como aditivo en las gasolinas sin plomo. Desde que está prohibida la utilización de los alquil- plomo debido a su toxicidad, el MTBE se añade a las gasolinas con el fin de aumentar su índice de octano. El índice de octano mide la "resistencia al cascabeleo o picado" de los carburantes cuando se queman mezclados con aire en la cámara de combustión de los motores. El MTBE se utiliza igualmente como compuesto oxigenado que permite aumentar el contenido en oxígeno de las gasolinas, y por esto, mejorar su eficacia de combustión. Esto permite reducir la expulsión a la atmósfera de los hidrocarburos no quemados, así como el de monóxido de carbono. Así, concentraciones de MTBE de 15% (v/v) se utilizan corrientemente en los carburantes oxigenados. El MTBE ha sido clasificado como compuesto cancerígeno potencial por la agencia de EE.UU. para la protección de medio ambiente (US E.P.A., Diciembre 1997 EPA/822/F-97/008. Office of Water, Washington, DC, USA). La contaminación por este compuesto del medio ambiente puede ser debida al almacenamiento inadecuado de la gasolina en los depósitos no estancos o a vertidos accidentales. Este tipo de residuos puede causar serios problemas de contaminación medioambiental, tal como la contaminación de los acuíferos subterráneos. Los consumidores pueden estar expuestos a débiles concentraciones cuando beben un agua no potable procedente de una fuente contaminada por el MTBE. Por otra parte, el gusto o sabor desagradables que el MTBE confiere al agua, incluso a débiles concentraciones, la hace inapropiada para el consumo, haciendo de este compuesto xenobiótico un contaminante importante. La contaminación por el MTBE parece ser principalmente atribuible al hecho de que sea eliminado difícilmente del medio ambiente. El MTBE se revela persistente por su elevada solubilidad en el agua y su débil biodegradabilidad. La semi-vida del MTBE en los acuíferos se estima en al menos 2 años (Wilson J. T., 2003, 19-61: In E. E. Moyer and P. T. Kostecki (ed.), *MTBE Remediation Handbook*. Amherst Scientific Publishers, Amherst, MA.); este valor se puede comparar con el del benceno que es de 2 a 3 meses en condiciones idénticas. La contaminación de los acuíferos puede engendrar serios riesgos para la salud pública. Por tanto es necesario desarrollar procedimientos eficaces que permitan tratar los acuíferos contaminados por el MTBE o cualquier otro intermedio catabólico del MTBE.

Desde hace algunos años se han emprendido varios estudios tendentes a determinar la biodegradabilidad del MTBE. Se han identificado y aislado cierto número de microorganismos capaces de asimilar completamente o parcialmente el MTBE. Por ejemplo, ciertos microorganismos son capaces de catabolizar el MTBE por co-metabolismo. Dos vías de degradación, que implican las dos la oxidación inicial del MTBE, conducen a la producción del alcohol *terc*.butílico (TBA) que, más generalmente se acumula en el medio. La cepa bacteriana *Mycobacterium vaccae* JOB5, cuando se cultiva sobre propano es capaz de oxidar el TBA aunque esto no permita generar compuestos útiles para su crecimiento. Se han podido aislar algunas cepas de microorganismos capaces de asimilar el MTBE como fuente de carbono y de energía (François et al., *Appl. Environ. Microbiol.*, 2002, **68**: 2754-2762; Hanson, J. R., et al. *Appl. Environ. Microbiol.*, 1999, **65**: 4788-4792 Hatzinger, P. B., et al., *Appl. Environ. Microbiol.*, 2001, **67**: 5601-5607). Todavía no han sido aclarados los mecanismos enzimáticos del ataque inicial del MTBE en estas bacterias.

La cepa bacteriana *Mycobacterium austroafricanum* I-2562 es una de las bacterias conocidas por su capacidad de crecer en condiciones aerobias, en presencia del MTBE. Esta cepa es en efecto capaz de catabolizar el MTBE en fuente de carbono y de energía (François et al., *Appl. Environ. Microbiol.*, 2002, **68**: 2754-2762; patente de EE. UU. nº 6.849.445). Se han identificado cierto número de intermedios de degradación o catabolitos del MTBE, tales como el TBA, el 2-M-1,2-PD, el hidroxibutiraldehído y el HIBA. La vía de degradación del MTBE para esta cepa bacteriana está descrita en la Figura 1.

La cinética de degradación del MTBE y del crecimiento que le acompaña en *M. austroafricanum* I-2562 presenta dos fases bien distintas (Figura 2). La cinética de degradación del MTBE se caracteriza por una velocidad de degradación muy rápida en el curso de las primeras 48 horas, que disminuye del día 2º al 16º. Durante esta etapa de transformación del MTBE en TBA, el TBA se acumula y por tanto no se observa ningún crecimiento. Del día 16º al 21º, el TBA acumulado se degrada induciendo al crecimiento del microorganismo.

Hay que advertir que la primera etapa de degradación comprende la mineralización del formiato en CO₂ lo que necesita que el microorganismo sea metilótrofo (François et al., *Appl. Environ. Microbiol.*, 2002, **68**: 2754-2762). La formiato- deshidrogenasa dependiente de NAD(P)⁺ (o FDH) es una enzima que desempeña un papel importante en la producción de energía en los microorganismos metilótrofos. Esta enzima, es por otra parte, frecuentemente utilizada en la regeneración del cofactor del tipo NAD(P)⁺ en las reacciones de biocatálisis (Tishkov, V. I. et al. *Biochemistr. (Moscú)*, 2004, **69**: 1537-1554). Esta primera etapa del catabolismo no genera ATP. François et al., han mostrado que el TBF tenía un efecto negativo sobre la velocidad de degradación del MTBE en *M. austroafricanum* I-2562 (François et al., *App Microbiol. Biotechnol.* 2003, **62**: 256-262). Se ha sugerido que el déficit en equivalentes reducidos de coenzimas y en energía era igualmente responsable de la ralentización de las reacciones enzimáticas (Salanitro, *Curr. Op. Biotechnol.*, 1995, **6**: 337-340). Estas dos razones son probablemente las responsables principales de la ralentización de la velocidad de degradación del MTBE.

La segunda etapa, al permitir la producción neta de 2 NAD(P)H y de 2H⁺, alimenta así la cadena de transporte de electrones con el O₂ como aceptor final de electrones en aerobiosis. Esta etapa parecería pues esencial en el metabolismo del MTBE. El resultado final de esta cadena de transporte de electrones es regenerar el NAD(P)⁺ a partir del NAD(P)H, reducir el O₂ a H₂O, creando un gradiente de transporte de protones que se utilizan para la formación de ATP por la ATP -sintasa (bomba de protones). Esto corresponde al mecanismo de fosforilación oxidante y permite la síntesis de ATP que es el compuesto energético de los microorganismos (30,5 kJ producto/mol de ATP degradado debido a un enlace de alto nivel energético).

Un objetivo de la invención es disponer de polipéptidos capaces de metabolizar el MTBE, y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente el TBA, el 2-M-1,2-PD, el hidroxiisobutiraldehído y el HIBA, y eventualmente modificarlos y utilizarlos con el fin de tratar efluentes contaminados por el MTBE, y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente el TBA, el 2-M-1,2-PD, el hidroxiisobutiraldehído y el HIBA.

Los trabajos de François et al., han permitido poner de manifiesto que varios polipéptidos eran específicamente inducidos en presencia de MTBE o TBA en *M. austroafricanum* IFP 2012 (I-2562). Estos polipéptidos específicamente inducidos han sido purificados, digeridos con tripsina y después microsecuenciados. El análisis de las secuencias resultantes con ayuda de las herramientas de alineamiento, los programas informáticos Blast y Fasta, ha sugerido que estos polipéptidos corresponden a óxido-reductasas implicadas en la degradación del MTBE y del TBA.

La firma solicitante ha realizado nuevas secuenciaciones peptídicas a partir de polipéptidos específicamente inducidos en presencia de MTBE, con el fin de obtener una primera sonda de 204 pares de bases. Sin embargo, esta sonda de 204 pares de bases se ha revelado demasiado corta para una hibridación sobre colonias. Por tanto se han realizado numerosos alineamientos por el programa BLAST con las proteínas más similares a la secuencia de aminoácidos deducida del fragmento de DNA de 204 pb. Después de la identificación de las secuencias de aminoácidos fuertemente conservadas en todas estas proteínas, se han realizado por PCR numerosas amplificaciones utilizando cebadores degenerados que permite obtener una nueva sonda de 604 pares de bases. Esta sonda de 604 pares de bases ha permitido realizar la clonación de un fragmento de DNA que comprende el gen mpdC, el orf1, el gen mpdB, el orf2 y una secuencia correspondiente a una supuesta transposasa. La clonación de este fragmento de DNA se ha revelado difícil, puesto que después de la una primera clonación, los clones obtenidos después de la transformación en *E. Coli* no eran estables, lo que hacía problemática la extracción del DNA plasmídico. Después de numerosos experimentos la firma solicitante ha emitido la hipótesis que la inestabilidad de los clones obtenidos procedía de la expresión del gen de la transposasa, y por tanto ha clonado un fragmento de DNA en el cual el gen de la transposasa está suprimido y que corresponde a la secuencia SEQ ID NO:11. Así, se han podido obtener clones estables que comprenden la secuencia SEQ ID NO:11.

Por tanto la firma solicitante ha debido enfrentarse a numerosas dificultades de clonación y secuenciación para obtener los polipéptidos de acuerdo con la invención.

La presente invención considera cualquier polipéptido aislado o purificado que tenga una actividad en la vía de degradación del MTBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente el TBA, el 2-M-1,2-PD, el hidroxiisobutiraldehído y el HIBA.

Al precisar que los polipéptidos o los ácidos nucleico que codifican al menos un polipéptido conforme a la invención son «aislados o purificados», se indica que son colocados en un ambiente diferente de que se encuentran naturalmente. Pueden ser aislados o purificados de una cepa bacteriana capaz de crecer en otro medio que comprende MTBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionados del grupo que consiste en el TBA, el 2-M-1,2-PD, el hidroxiisobutiraldehído y el HIBA, siendo algunos de ellas principalmente las cepas *Mycobacterium austroafricanum* registradas en la colección CNCM bajo los números I-3401

y I-2532. Se entienden igualmente que las moléculas que han sido alteradas por la presente invención están «aisladas». Se recuerda que estos calificativos «aislado o purificado» se utilizan igualmente para las células hospedantes.

5 En el contexto de la invención, los términos «polipéptido, enzima» y «proteína» pueden ser utilizados de manera intercambiable y designan moléculas caracterizadas por secuencias de aminoácidos de cualquier longitud, eventualmente modificadas por vía química o bioquímica. Igualmente se incluye en el término «polipéptido» el conjunto de polipéptidos mutados que pueden existir naturalmente o variantes, en particular en la bacteria de la cepa *M. austroafricanum*, y que corresponden a sustituciones, deleciones, inserciones o adiciones de al menos un aminoácido. En el caso de una sustitución, uno o varios aminoácidos consecutivos o no consecutivos, pueden estar
10 reemplazados por aminoácidos «equivalentes». La expresión aminoácido «equivalente» pretende aquí designar cualquier aminoácido susceptible de ser sustituido por uno de los aminoácidos de la estructura de base, sin modificar, no obstante, la actividad biológica de los polipéptidos según la invención. Estos aminoácidos equivalentes pueden ser determinados ya sea apoyándose en su homología de estructura con los aminoácidos a los cuales sustituyen, o bien sea sobre los resultados de los ensayos de actividad biológica cruzada a los cuales los diferentes
15 polipéptidos son susceptibles de dar lugar.

El polipéptido según la invención se selecciona del grupo que consiste en:

- a) un polipéptido que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en la SEQ ID NO:2 o la SEQ ID NO:10,
- 20 b) un polipéptido que comprende una secuencia de aminoácidos que presente al menos 70 % de identidad, preferentemente 75%, 80%, 90%, 95%, 98% o 99% de identidad con la secuencia de aminoácidos de un polipéptido que comprende la SEQ ID NO:2, el polipéptido tal como el definido en a) o b) de la secuencia SEQ ID NO:2 que tiene una actividad de deshidrogenación del hidroxibutiraldehído a HIBA, y

el polipéptido tal como el definido en a) de secuencia SEQ ID NO:10 que tiene una actividad de regulador transcripcional.

25 Se han descrito diferentes protocolos conocidos por los expertos en la técnica para introducir mutaciones en los polipéptidos. Típicamente para modificar una secuencia de aminoácidos de un polipéptido, se actúa sobre la molécula de ácido nucleico que codifica dicho polipéptido. Es posible para modificar el ácido nucleico que codifica un polipéptido conforme a la invención, someterlo a un tratamiento por un agente mutágeno, es decir, un agente físico o químico capaz de provocar mutaciones que alteren el significado de los codones, lo que por el juego del código genético modifica la secuencia de aminoácidos. Ventajosamente, se recurre a una técnica de mutagénesis dirigida para modificar la secuencia nucleotídica de un ácido nucleico que codifica al menos un polipéptido del gen, así como las de polinucleótidos conformes a la invención; se trata entonces de introducir específicamente una o varias mutaciones en el ácido nucleico estudiado, lo que por el juego del código genético implica la sustitución de uno o varios aminoácidos por uno o varios otros aminoácidos en el polipéptido codificado por el polinucleótido mutado. El
30 interés de realizar estas mutaciones es no solamente estudiar la actividad biológica de uno de los polipéptidos conformes a la invención en la vía de degradación del MTBE y/o al menos uno de los intermedios catalíticos, sino eventualmente optimizar dicha actividad del polipéptido bajo forma recombinante en vista de su aplicación industrial.

Por «fragmento de polipéptido», se designa un polipéptido que comprende como mínimo 15 aminoácidos consecutivos, de preferencia 17, 20, 23, 25, 30, 40, 50, 100, 250 o 300 aminoácidos consecutivos. Los fragmentos
40 de polipéptido conforme a la invención se pueden obtener por escisión de dicho polipéptido por una enzima proteolítica, por un reactivo químico, o incluso al colocar dicho polipéptido en un ambiente muy ácido forma igualmente parte de la invención.

Por «fragmento biológicamente activo», se entiende un fragmento de secuencias de aminoácidos de un polipéptido conforme a la invención que tiene al menos una de las características o propiedades funcionales de dicho polipéptido, principalmente una actividad en la vía de degradación del MTBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente el TBA, 2-M-1,2-PD, el hidroxibutiraldehído y el HIBA.

Según un modo preferido de realización, la expresión de los polipéptidos conformes a la invención es inducida cuando las bacterias susceptibles de asimilar el MTBE y/o al menos uno de los intermedios catabólicos del MTBE, de las cuales principalmente las bacterias de la cepa *M. austroafricanum*, más preferentemente las bacterias I-2562
50 o I-3401, están en un medio que comprende el MTBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente el TBA, el 2-M-1,2-PD, el hidroxibutiraldehído y el HIBA. Preferentemente, los polipéptidos cuya expresión es fuertemente inducida son los polipéptidos de la SEQ ID NO:2, tal como se ha definido en a) o b), de las SEQ ID NO:4 y SEQ ID NO:6, como se indica en la Figura 3.

Ventajosamente, el polipéptido de la SEQ ID NO:2 tal como se ha definido en a) o b) denominado MpdC tiene una actividad de aldehído-deshidrogenasa. En la vía catabólica del MTBE, el polipéptido MpdC conforme a la invención es preferentemente susceptible de deshidrogenar el hidroxibutiraldehído a HIBA.

Ventajosamente, el polipéptido de la SEQ ID NO:6 denominado MpdB tiene una actividad de alcohol-deshidrogenasa. En la vía catabólica del MTBE, el polipéptido MpdB conforme a la invención es preferentemente susceptible de deshidrogenar el 2-M-1,2-PD a hidroxisobutiraldehído.

5 Ventajosamente, el polipéptido de la SEQ ID NO:8 tiene una actividad de permeasa, preferentemente de di/tripéptido-permeasa. Dicho polipéptido comprende preferentemente 5 segmentos transmembranales.

Ventajosamente, el polipéptido de la SEQ ID NO:10 tal como se ha definido en a) y denominado MpdR tiene una actividad de regulador transcripcional.

10 La presente invención se refiere igualmente a cualquier ácido nucleico purificado o aislado que codifica al menos un polipéptido conforme a la invención. Preferentemente, la presente invención tiene por objeto cualquier ácido nucleico aislado o purificado que codifica al menos un polipéptido que tiene una actividad en la vía de degradación del MTBE, o al menos uno de los intermedios catabólicos del MTBE, preferentemente el TBA, el 2-M-1,2-PD, el hidroxisobutiraldehído y el HIBA.

15 En el contexto de la invención, los términos «ácido nucleico» y «DNA» se utilizan de manera intercambiable. Por «ácido nucleico», se designa un encadenamiento preciso de nucleótidos, modificados o no, que permiten definir un fragmento o una región de un ácido nucleico, que comprende o no nucleótidos no naturales, y que pueden corresponder también a una DNA bicatenario, un DNA monocatenario, así como a productos de transcripción de dichos DNA, y/o un fragmento de RNA.

El ácido nucleico según la invención se selecciona del grupo que consiste en:

20 e) un ácido nucleico que comprende al menos una cualquiera de las secuencias nucleotídicas seleccionada del grupo que consiste en la SEQ ID NO:1 o SEQ ID NO:9, o un complemento de éstas,

f) un ácido nucleico que comprende al menos una secuencia nucleotídica que presenta al menos 70 % de identidad, preferentemente 75%, 80%, 90%, 95%, 98% o 99% de identidad con la secuencia nucleotídica de un ácido nucleico tal como se ha definido en e).

25 Por «complemento» se designa cualquier ácido nucleico cuyos nucleótidos son complementarios de los de la secuencia SEQ ID NO:1 o SEQ ID NO:9, y cuya orientación está invertida.

30 Por «porcentaje de identidad» entre dos secuencias de ácidos nucleicos o de aminoácidos en el sentido de la presente invención, se designa un porcentaje de nucleótidos o de aminoácidos idéntico entre las dos secuencias que se comparan. El porcentaje de identidad entre dos secuencias de ácidos nucleicos o de aminoácidos se determina comparando estas dos secuencias alineadas de manera significativa, pudiendo comprender la secuencia de ácidos nucleicos o de aminoácidos que se comparan adiciones o deleciones con relación a la secuencia de referencia para un alineamiento significativo entre estas dos secuencias. El porcentaje de identidad se calcula determinando el número de posiciones idénticas para las cuales el nucleótido o el aminoácido es idéntico entre las dos secuencias, dividiendo este número de de posiciones idénticas por el número total de posiciones comparadas y multiplicando el resultado obtenido por 100 para obtener el porcentaje de identidad entre estas dos secuencias. Por «alineamiento significativo», se designa el alineamiento para el cual el porcentaje de identidad determinado como se ha indicado es el más elevado. Preferentemente, se utilizará el programa informático BLAST para obtener un alineamiento significativo.

40 Ventajosamente, el ácido nucleico conforme a la invención, comprende las secuencias nucleotídicas SEQ ID NO:1 tal como se ha definido en e) o f) y la SEQ ID NO:5, preferentemente la SEQ ID NO:1 tal como se ha definido en e) o f), la SEQ ID NO:3 y la SEQ ID NO:5.

Ventajosamente, el ácido nucleico conforme a la invención comprende la secuencia nucleotídica SEQ ID NO:11.

45 Según un modo preferido de realización, la transcripción del ácido nucleico conforme a la invención está bajo el control de un promotor único. Por «promotor», se designa una región de regulación situada en dirección 5' de un marco de lectura abierto (ORF), en la proximidad de su extremo 5'. En la presente solicitud, se utilizarán los términos «ORF» y «gen» de manera intercambiable. Un promotor comprende ciertas secuencias nucleotídicas características que permiten la fijación del complejo de iniciación de la transcripción así como la fijación de reguladores transcripcionales.

50 Según otro modo preferido de realización, el ácido nucleico conforme a la invención se caracteriza porque su transcripción es inducida por la presencia de MTBE, y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionados del grupo que consiste en el TBA, el 2-M-1,2-PD, el hidroxisobutiraldehído y el HIBA. Ventajosamente, el ácido nucleico conforme a la invención está organizado en operón.

Según un modo preferido de realización, el ácido nucleico conforme a la invención corresponde a un grupo que comprende los ácidos nucleicos que codifican las SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7 y SEQ ID NO:9, siendo dicho grupo denominado grupo *mpd*. El grupo *mpd* está esquematizado en la Figura 7.

Según un modo preferido de realización, el ácido nucleico conforme a la invención es un ácido nucleico recombinante. Por «ácido nucleico recombinante», se designa una molécula de ácido nucleico, monocatenario o bicatenario, que ha sido modificada por una intervención humana de manera que contenga fragmentos de ácidos nucleicos combinados o yuxtapuestos según una disposición no existente en el estado natural.

- 5 La invención tiene igualmente por objeto cualquier vector de expresión y/o de clonación que comprenda un ácido nucleico conforme a la invención. Por «vector» se designa una molécula de ácido nucleico extracromosómico, principalmente un plásmido, que tenga una replicación autónoma y que pueda ser incorporada en una «célula hospedante».

10 Típicamente, se concibe un vector de clonación para permitir el transporte de un fragmento de DNA clonado y contiene uno o varios sitios de reconocimiento por enzimas de restricción que permiten la inserción o clonación de un fragmento de ácido nucleico, así como una o varias secuencias nucleotídicas codificantes para genes que permiten la identificación y la selección de células hospedantes transformadas por dicho vector, tales como genes de resistencia a antibióticos.

15 Según un modo preferido de realización, el vector de expresión conforme a la invención es concebido para permitir la expresión de una secuencia nucleotídica codificante insertada en dirección 3' de un promotor. La secuencia insertada o inserto es entonces transcrita, y después traducida a polipéptido. Ventajosamente, el vector conforme a la invención está flanqueado por elementos que aseguran la expresión de al menos un ácido nucleico conforme a la invención en la célula hospedante. Ventajosamente ciertos vectores de expresión tienen una secuencia que codifica una etiqueta en dirección 5' o en dirección 3' del sitio de inserción o de clonación; el ácido nucleico insertado es entonces transcrito y después traducido bajo la forma de una proteína de fusión. Por «proteína de fusión» se designa un polipéptido híbrido que comprende un polipéptido codificado por un ácido nucleico conforme a la invención y un polipéptido capaz de fijarse a matrices de afinidad y/o de ser reconocido por anticuerpos que permiten detectar y/o de purificar el polipéptido al cual está fusionado.

20

25 Según un otro modo preferido de realización, el vector conforme a la invención comprende al menos una secuencia homóloga de una secuencia de ácido nucleico presente en el genoma de una célula hospedante y que asegura su integración en dicho genoma.

La invención se refiere además a una célula hospedante aislada que comprende al menos uno de los ácidos nucleicos conforme a la invención y/o al menos uno de los vectores conforme a la invención. Entre las células hospedantes utilizables en el sentido de la presente invención, se puede citar una célula procariota, preferentemente una célula bacteriana, más preferentemente una célula bacteriana seleccionada del grupo que consiste en *Mycobacterium smegmatis* mc2 155, *Escherichia Coli*, *Rhodococcus ruber* (registrada en la colección CNCM bajo el nombre *Gordonia terrae* y el número CIP I-1885).

30

Los polipéptidos conformes a la invención pueden ser preparados y/o obtenidos por cualesquiera técnicas dominadas por los expertos. Dichos polipéptidos se pueden obtener principalmente por síntesis química pero igualmente por técnicas de biología molecular, utilizando principalmente la PCR, vectores de expresión y células hospedantes apropiadas tales como las descritas anteriormente. Preferentemente, los polipéptidos conformes a la invención se caracterizan porque su expresión en células hospedantes, principalmente células bacterianas, permiten satisfacer la exigencia de degradación del MTBE, y/o al menos uno de los intermedios catabólicos y permiten el crecimiento de dichas bacterias.

35

40 La secuencia SEQ ID NO:12 corresponde a la secuencia nucleotídica del rDNA 16S de la cepa *M. austroafricanum* I-2532.

Par «sonda» se entiende un fragmento de ácido nucleico marcado por la incorporación de átomos radiactivos o de grupos fluorescentes y cuya secuencia es sensiblemente complementaria a la secuencia de un ácido nucleico buscado; este último será descubierto/detectado por hibridación con la sonda, produciéndose dicha hibridación cuando se aparean las dos secuencias complementarias. Por «secuencia complementaria» se designa una secuencia nucleotídica constituida por una sucesión de bases complementarias a la de otra secuencia con la cual es por tanto capaz de hibridarse. Preferentemente, la hibridación de una sonda cuyo tamaño sea superior a 200 nucleótidos se realiza a una temperatura de aproximadamente 60°C según la enseñanza adaptada de Sambrook et al., (*Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 2001).

45

50

La presente invención se refiere igualmente a un procedimiento de identificación de una célula o de un ácido nucleico de una célula susceptible de degradar el MTBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionado del grupo que consiste en el TBA, el 2-M-1,2-PD, el hidroxisobutiraldehído y el HIBA, que comprende:

- 55 - opcionalmente una etapa de siembra de la célula en un medio suplementado en MTBE, y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionado del grupo que consiste en el TBA, el 2-M-1,2-PD, el hidroxisobutiraldehído y el HIBA,

- una etapa de cribado de los ácidos nucleicos de dicha célula por hibridación con al menos una sonda, comprendiendo dicha sonda la secuencia nucleotídica complementaria de un fragmento de al menos 50 nucleótidos consecutivos de la secuencia nucleotídica seleccionada del grupo que consiste en la SEQ ID NO:1 o SEQ ID NO:9, o comprendiendo dicha sonda un fragmento del gen mpdB de 591 pares de bases que tiene la secuencia SEQ ID NO: 14.

Según un modo preferido de realización, una célula de una nueva cepa capaz de degradar el MTBE, y/o al menos uno de los intermedios catabólicos del MTBE, será identificada gracias a cebadores oligonucleotídicos que permitan la amplificación del rRNA 16S de dicha célula.

Ventajosamente, se ha concebido una pareja de cebadores (5'-TGACACAGGCCACAACCCA-3') y (5'-GAGAGTTTGATCCTGGCTCAG-3') a partir de dos regiones variables de la secuencia nucleotídica del rDNA 16S de la cepa *M. austroafricanum*. Esta pareja de cebadores es específica de la especie *M. austroafricanum* en la medida en la que no permite amplificar el rDNA 16S de microorganismos de otras especies, de las cuales principalmente la especie de las *Nocardiacae*, incapaces de asimilar el MTBE, y/o al menos uno de los intermedios catabólicos del MTBE (Figura 12).

Empleando el procedimiento conforme a la invención se ha identificado un nueva cepa bacteriana *M. austroafricanum*; está registrada en la colección CNCM bajo el número I-3401. Ha sido aislada del agua de chorreo del prelavado del fondo de un depósito de almacenamiento de una gasolina aditivada con MTBE. El crecimiento de *M. austroafricanum* I-3401 en MTBE sigue una vía de degradación similar a la de *M. austroafricanum* I-2562 (Figura 10). Preferentemente, la nueva cepa *M. austroafricanum* I-3401 se caracteriza porque sea capaz de asimilar el MTBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionados en el grupo que consiste de el TBA, el 2-M-1,2-PD, el hidroxisisobutiraldehído y el HIBA. Más preferentemente, la nueva cepa I-3401 se caracteriza por una secuencia nucleica del rDNA 16S que comprende la secuencia SEQ ID NO:13.

El aislamiento independiente de dos cepas de *M. austroafricanum*, I-2562 y I-3401, ambas dos capaces de crecer en MTBE y aisladas de dos orígenes geográficos distintos y alejados (lodos activados de una estación de depuración de una factoría de tratamiento de aguas urbanas de la región parisiense y agua de chorreo del prelavado del fondo de un depósito de almacenamiento de una gasolina aditivada con MTBE localizada en el sudoeste de Francia, respectivamente) muestra que tales microorganismos, que se aíslan frecuentemente de muestras de suelos y de aguas (Jones y Jenkins, *Can. J. Microbiol.*, 1965, **11**: 127-133; Viallier y Viallier, *Ann. Soc. Belge Med. Trop.*, 1973, **53**: 361-371), pueden pues desempeñar un papel significativo en la degradación del MTBE en los acuíferos contaminados.

La presente invención se refiere igualmente a la utilización del polipéptido SEQ ID NO:2 tal como se ha definido en a) o b) en calidad de aldehído-deshidrogenasa, preferentemente para la deshidrogenación del hidroxisisobutiraldehído a HIBA.

La presente invención se refiere finalmente a un procedimiento de tratamiento de efluentes acuosos que comprenden del MTBE, el ETBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionado del grupo que consiste en el TBA, el 2-M-1,2-PD, el hidroxisisobutiraldehído y el HIBA, con el fin de reducir su concentración, que comprende la siembra de dicho efluente acuoso con al menos una célula conforme a la invención o identificada por el procedimiento de identificación conforme a la invención. Ventajosamente, el procedimiento conforme a la invención se caracteriza porque la célula es fijada sobre un soporte adaptado, preferentemente un soporte mineral, más preferentemente un soporte que comprende al menos 50% en masa de partita.

El procedimiento según la invención está particularmente adaptado para tratar un acuífero o un suelo contaminado por del MTBE y/o al menos uno de los intermedios catabólicos del MTBE.

LISTA DE LAS FIGURAS DE LOS DIBUJOS

Figura 1. Vía metabólica de degradación del MTBE por *M. austroafricanum* I-2562. En grueso, los intermedios de degradación que han sido identificados en el curso de nuestras experimentaciones.

Figura 2. Crecimiento de *M. austroafricanum* I-2562 sobre MTBE. Durante el crecimiento se midieron la D.O. a 600 nm (\diamond - \diamond) así como las concentraciones en MTBE (\blacksquare - \blacksquare), TB (Δ - Δ) y TBA (-).

Figura 3. Electroforesis sobre geles de poliacrilamida al 10% (A) o 15% (B) en presencia de SDS (SDS-PAGE) de los extractos celulares de *M. austroafricanum* I-2562 obtenidos después de crecimiento sobre glucosa o sobre MTBE. Los pesos moleculares (en kDa) están indicados a la derecha de la figura.

Figura 4. Amplificación por PCR sobre DNA genómico de *M. austroafricanum* I-2562 con los cebadores degenerados creados a partir de las secuencias peptídicas internas del polipéptido de 64 kDa inducido en el MTBE.

Figura 5. Hibridación por transferencia Southern del DNA genómico de *M. austroafricanum* I-2562 después de digestión utilizando la sonda de 204 pb. El tamaño de las bandas reveladas está indicado en kb. 5 μ g de DNA

genómico de *M. austroafricanum* I-2562 son digeridos por los enzimas de restricción *Bam*HI, *Eco*RI, *Fsp*I, *Kn*pI, *Nhe*I, *Nru*I, *Pst*I, *Pvu* II o *Sma*I (pistas 2-8, respectivamente) y 50 ng de sonda de 204 pb (pista 12) son analizados por Southern Blot utilizando la sonda de 204 pb. Las posiciones correspondientes a la migración del DNA del marcador de peso molecular III marcado con digoxigenina (Roche; pista 1 y pista 9) son indicadas a la izquierda de la figura.

5 **Figura 6.** "Hibridación sobre colonias" de los clones recombinantes de *E. coli* DH10B con una sonda de 604 pb.

A) Luminografía de las membranas que contienen los primeros clones que proceden de la digestión por *Sma*I de *M. austroafricanum* I-2562. Los clones positivos (manchas negras) son detectados con la sonda *mpdC* marcada con digoxigenina.

10 B) Luminografía de las membranas que contienen los segundos clones procedentes de la digestión por *Pst*I de *M. austroafricanum* I-2562. Los clones positivos (manchas negras) son detectados con la sonda *mpdB* marcada con digoxigenina.

Figura 7. Organización del grupo de los genes implicados en la vía de degradación del MTBE y aislado a partir del DNA genómico de *M. austroafricanum* I-2562 (A) y su mapa de restricción correspondiente (B).

15 El ORF *mpdR* codifica un regulador transcripcional, los ORF *mpdC* y *mpdB* codifican respectivamente una enzima aldehído-deshidrogenasa y una enzima alcohol-deshidrogenasa. Tres polipéptidos de 64, 55 y 27 kDa son todos fuertemente expresados después del crecimiento de *M. austroafricanum* I-2562 sobre MTBE. El polipéptido de 25 kDa es una permeasa. Los fragmentos de PCR procedentes de las amplificaciones realizadas con ayuda de parejas de cebadores RT-PCR-F1/-R1, RT-PCR-F2/-R2 y RT-PCR-F3/-R3 están representados esquemáticamente en la Figura 7 (fragmentos numerados I, II y III).

20 **Figura 8.** RT-PCR sobre *M. austroafricanum* I-2562 después del crecimiento sobre TBA utilizando los cebadores correspondientes al grupo *mpd*.

Los cebadores utilizados se citan en la Tabla que se encuentra más adelante.

25 Leyenda: Pista 1: C⁻ PCR I, Pista 2: C⁺ PCR I, Pista 3:(-) RT-PCR I, Pista 4: (+) RT-PCR I, Pista 5: C⁺ PCR II, Pista 6: (-) RT-PCR II, Pista 7: (+) RT-PCR II, Pista 8: C⁺ PCR III, Pista 9: (-) RT-PCR III, Pista 10: (+) RT-PCR III, Pista 11: marcador de DNA (100 pb).

Figura 9. Cinética de degradación del 2-M-1,2-PD por células de *M. smegmatis* mc2 155-clon 9 (cuadrados) y de *M. smegmatis* mc2 155-pCL4D (triángulos).

Degradación del 2-M-1,2-PD (símbolos rellenos) y producción del HIBA (símbolos vacíos). Control abiótico (◇-◇). Células de *M. smegmatis* mc2 155-clon 9 en presencia de 400 mg.L⁻¹ de cloranfenicol (x-x).

30 Las concentraciones celulares de *M. smegmatis* mc2 155-clon 9 y de *M. smegmatis* mc2 155-pCL4D eran 233,33 ± 14,09 y 214,63 ± 21,95 mg.L⁻¹, respectivamente.

Figura 10. Crecimiento de *M. austroafricanum* I-3401 sobre MTBE.

Durante el crecimiento, se midieron la D.O._{600nm} (◆-◆), y las concentraciones residuales en MTBE (□-□), TBF (Δ-Δ) y TBA (▲-▲).

35 **Figura 11.** Amplificación por PCR del grupo *mpd* en el DNA genómico e de *M. austroafricanum* I-3401.

Leyenda:

MM: marcador de DNA 1 kb,

Pistas 1 y 6: amplificación de SEQ ID NO:9 que codifica el polipéptido de 47 kDa.

Pistas 2 y 7: amplificación de SEQ ID NO:1 que codifica la aldehído-deshidrogenasa de 55 kDa.

40 Pistas 3 y 8: amplificación de SEQ ID NO:3 que codifica la proteína de 27 kDa

Pistas 4 y 9: amplificación de SEQ ID NO:5 que codifica la alcohol-deshidrogenasa de 64 kDa.

Pistas 5 y 10: amplificación de SEQ ID NO:7 que codifica la proteína de 25 kDa, la supuesta permeasa.

C 1 y C2: controles negativos sin cebadores de los DNA genómicos de *M. austroafricanum* I-2562 y I-3401, respectivamente.

45 **Figura 12.** Amplificación de secuencias específicas del rDNA 16S en diversos microorganismos.

El tamaño esperado del producto de PCR es 331 pb.

Los DNA genómicos de las cepas de *Rhodococcus ruber* (Pista 1), *Rhodococcus* sp. B-1 (Pista 2), *Pseudomonas resinovorans* CA10 (Pista 3), *Mycobacterium smegmatis* mc2 155 (Pista 4), *Escherichia coli* DH10B (Pista 5), *Mycobacterium austroafricanum* I-2562 (Pista 6), *Mycobacterium austroafricanum* I-3401 (Pista 7) y muestras que proceden del medio de la casete 1 (Pista 9) y del medio de la casete 2 (Pista 10) de una biobarrera piloto sembrada con *M. austroafricanum* I-2562 para la degradación del MTBE se utilizaron para efectuar las reacciones de amplificación por PCR. Como marcador de pesos moleculares del DNA (Pista 8) se utilizó *GeneRuler™* DNA escalera de 1 kb.

La presente invención será mejor comprendida con ayuda de los ejemplos que siguen; estos ejemplos se dan únicamente a modo de ilustración de los objetos de la invención, y por tanto de ningún modo constituyen una limitación.

EJEMPLO 1: IDENTIFICACIÓN DE LOS POLIPÉPTIDOS INDUCIDOS DESPUÉS DEL CRECIMIENTO DE LA CEPA *M. AUSTROAFRICANUM* I-2562 EN MTBE

Con el fin de identificar y de detectar la presencia de proteínas inducidas específicamente después del crecimiento de las bacterias *M. austroafricanum* en un medio que comprende el MTBE, se han obtenido extractos de proteínas citoplásmicas a partir de cultivo en MTBE o en glucosa (Figura 2).

1.1. Preparación de los extractos proteicos

Células de la cepa *M. austroafricanum* I-2562 se cultivan en 300 mL de medio mineral (MM) definido precedentemente descrito (Piveteau et al., *Appl. Microbiol. Biotechnol.* 2001, **55**: 369-373) en presencia de MTBE o de glucosa como fuente de carbono. Los cultivos se incuban en condiciones aerobias a 30°C en frascos cónicos bajo agitación. El crecimiento se evalúa por medida de la absorbancia (D.O._{600nm}) en un espectrofotómetro UV-1601 (Shimadzu Corporación, Kyoto, Japón). Cuando la D.O.₆₀₀ es 1, las células son recogidas por centrifugación a 20,000 x g durante 15 minutos, se lavan dos veces en tampón de fosfato tampón (20 mM, pH 7) y se vuelven a poner en suspensión en 5 mL de tampón Tris-HCl a 50 mM (pH=8,0) que contiene de ditiotreitól (DTT) 0,1M. Las células son atenuadas por tres pases en la prensa French (20.000 psi = 137,895 MPa) y mantenidas siempre en hielo. Los residuos celulares se eliminan por dos centrifugaciones a 1.000 x g durante 2 minutos. El líquido sobrenadante se utiliza luego para análisis del perfil proteico. La concentración en proteínas totales se mide por el kit Bio-Rad (*Dye reagent*, Bio-Rad Laboratories GmbH, Munich, Alemania).

1.2. Análisis de las proteínas solubles

La electroforesis de estos extractos se efectúa en geles de SDS-PAGE con 10 o 15% de poli(acrilamida) que contiene SDS según el método puesto a punto por Laemmli empleado en parte para las condiciones de migración (V =150 voltios durante la primera hora y después A=33 mA de modo constante durante las horas siguientes de la migración). 12,5 µg de proteínas totales se depositan por pista en el gel, para cada fuente de carbono ensayada (Laemmli, *Nature*, 1970, **227**: 680-685).

Los perfiles proteicos de altos y bajos pesos moleculares revelados en los geles de SDS-PAGE con 10 o 15% de acrilamida, respectivamente, han permitido detectar 9 proteínas de pesos moleculares diferentes especialmente inducidas por la presencia del MTBE (Figura 3).

1.3. Identificación de dos secuencias peptídicas del polipéptido MpdC (64 kDa) inducido durante el crecimiento en MTBE

Protocolo del secuenciación interna:

Cada banda de gel correspondiente a un polipéptido inducido específicamente en el perfil proteico después del crecimiento en MTBE de *M. austroafricanum* I-2562 se recorta y después se somete a una reducción por DTT "en el gel" y después a una alquilación de los puentes disulfuro por yodoacetamida (Jenö et al., *Analytical Biochemistry*. 1995, **224**: 75-82) antes de la digestión con tripsina. La enzima utilizada es la tripsina de calidad "de secuenciación de Promega (Helman et al., 1995, William et al., *In: Techniques VIII, Marshak, D., ed. Academic Press, San Diego*, 1997, 79-90). Los péptidos así generados son entonces extractos de trozos de gel residuales por 3 tratamientos de 30 minutos a 60°C, cada uno en presencia de 100 µL de TFA al 1% y acetonitrilo al 60% seguido de un pase por el aparato de ultrasonidos durante 10 minutos. Una última extracción se efectúa en 50 µL de acetonitrilo puro durante 10 minutos. Los 4 extractos se mezclan y el volumen líquido se reduce en un evaporador rotatorio (Savant AES 1010) para obtener un volumen final de 5 µL.

Los péptidos se separan entonces por HPLC (Vydac) en una columna *Microbore* C18 (300Å; 1 X 50mm Nº 218TP5105) utilizando el sistema de separación de Applied Biosystems 130A. Los péptidos se eluyen a 100 µL/min por una mezcla de de disolventes A y B con un gradiente de mezcla de los disolventes programado así: 3-63 min (0-50% B), 63-72 min (50-100% B) y 72-75 min (100% B). El disolvente A está constituido por TFA al 0,1% en el agua y el disolvente B está compuesto por TFA al 0,08 % en 70% de acetonitrilo y agua, y los péptidos son detectados por

un detector UV a 220 nm. Las diferentes fracciones se recogen manualmente en tubos de 1,5 mL y se depositan en discos de fibra de vidrio pretratados por adición de 7 µL de biobreno (Applied Biosystems Inc.). Los discos se someten a la degradación de Edman en un secuenciador Procise (modelo 494 cLC) según el protocolo descrito por Hewick et al., (*J. Biol. Chem.*, 1981, **256**:7990-7997).

- 5 Una cantidad de péptido equivalente a aproximadamente 1 pmol se deposita en el secuenciador y se utiliza para la secuenciación un programa estándar que utiliza TFA como fase líquida. Los derivados aminoácidos-feniltiohidantoína (PTH-aa) se determinan por comparación con patrones (*PTH-standards*, ABI) y análisis en línea gracias a un sistema de separación capilar (ABI 140D) desde el arranque de la secuenciación.

Resultados:

- 10 Entre las proteínas identificadas, una proteína que migra de 64 kDa (denominada MpdC) es la más fuertemente inducida. La extracción a partir de la banda de gel correspondiente, la digestión con tripsina seguida de micro-secuenciaciones peptídicas internas ha permitido la obtención de 2 secuencias de este polipéptido.

- 15 La comparación con ayuda del programa informático BLAST de la primera secuencia interna (**KQRGWAYDPNVRGLPE**) no reveló similitudes significativas. Por el contrario el análisis por BLAST de la segunda secuencia interna (**STHEGLEGTDWPISYEELAPYYDENDAIY**) muestra una fuerte similitud (93%) con numerosas oxidasas y deshidrogenasas que pertenecen a la familia de las GMC-óxido-reductasas (Cavener, *J. Mol. Biol.*, 1992, **223**: 811-814).

EJEMPLO 2: PREPARACIÓN DE LAS SONDAS

2.1. Extracción del DNA genómico

- 20 El DNA genómico se aisló de las diferentes cepas ensayadas según el protocolo de Pospiech y Neumann modificado de la manera siguiente: un cultivo efectuado en 50 mL de medio LB (caldo de Luria) (D.O. a 600 nm de 0,7) se recoge por centrifugación a 10000 g durante 15 minutos a 4°C (Pospiech and Neumann, *Trends In Genetics.*, 1995, **11**: 217-218). El sedimento se vuelve a poner en suspensión en 5 mL de tampón SET (NaCl 75 mM, EDTA 25 mM, Tris-HCl 20 mM a pH 8). Se añade lisozima a una concentración final de 1 mg.mL⁻¹ así como lisostafina de 112 U.µL⁻¹ (Sigma). Esta mezcla de reacción se incuba durante 1 h a 37°C. Después de tratamiento por la mezcla SDS/Proteinasas K, se efectúa una etapa de extracción en la mezcla de reacción dependiendo de donde se añaden secuencialmente los diferentes reactivos: ¼ de volumen de NaCl 5 M seguido de 1/7,5 volumen de CTAB al 10% - NaCl 0,7 M (pre-calentado a 65°C) y la mezcla se incuba durante 1 hora a 65°C. Se añade a continuación un volumen de cloroformo (CHCl₃) y las etapas subsiguientes se realizan según el protocolo original.

- 30 2.2. Preparación de una sonda de 204 pb

- Las dos secuencias peptídicas obtenidas en el EJEMPLO 1 se utilizaron para crear varios pares de cebadores degenerados de acuerdo con el código genético universal y teniendo en cuenta dos posibilidades de apareamiento de las secuencias en el seno de la secuencia polipeptídica. El conocimiento de estas secuencias peptídicas permite, por el código genético, deducir las secuencias nucleotídicas correspondientes. El código genético equivale a la correspondencia entre las secuencias nucleotídicas y las secuencias peptídicas. Así, cada triplete de nucleótidos de la secuencia nucleotídica o codón corresponde a un aminoácido único. Se habla de degeneración del código genético en la medida en la que varios codones sinónimos pueden especificar un mismo aminoácido. En consecuencia, por el código genético, varias secuencias nucleotídicas pueden ser deducidas para una secuencia peptídica única. A partir de esta informaciones, se conciben y sintetizan cebadores oligonucleotídicos degenerados: dichos cebadores sirven como tales en el cuadro de una reacción de amplificación como resultado de la cual se obtiene una sonda que tiene una gran especificidad para el o los ácidos nucleicos buscados.

Una sola pareja de cebadores (MadFI/MadRI) permitió la obtención por PCR, a partir de DNA genómico puro de *M. austroafricanum* I-2562, de un fragmento de ácido nucleico de 204 pb (Figura 4).

La secuencia nucleotídica del cebador MadF1 es:

- 45 5'-GGNTGGGCNTAYGAYCC-3'

La secuencia nucleotídica del cebador MadRI es:

5'-GCRTCRTTYTCRTCSTAST-3'

La secuencia de aminoácidos se obtuvo gracias a un programa informático denominado *ORF finder*, véase la web:

(<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). Dicha secuencia está representada:

- 50 **KQRGWAYDPNVRGLPE DYPVTGFTTPYLMNNVGGSTMHYAGHWPRYKPVDFRK
GTEHGLEGTIDWPISYEELAPYYDEND.**

El análisis por el programa BLAST de esta secuencia de aminoácidos confirma la pertenencia de la proteína MpdC a la familia de las GMC-óxido-reductasas puesto que el análisis muestra 62 % de similitud (44% de identidad en 45 aminoácidos) con una supuesta colina-deshidrogenasa de *Bradyrhizobium japonicum*, 53 % de similitud (38% de identidad en 76 aminoácidos) con una glucosa-deshidrogenasa de una cepa de *Burkholderia cepacia* y 61 % de similitud (38% de identidad en 47 aminoácidos) con una óxido-reductasa, todas pertenecientes a esta misma familia.

2.3. Preparación de una sonda de 604 pb

Como consecuencia de las informaciones obtenidas a partir de esta primera amplificación por PCR, se concibieron y sintetizaron cebadores degenerados teniendo en cuenta ciertos restos conservados que están presentes en las proteínas de la familia de las GMC-óxido-reductasas. Una de estas parejas de cebadores (MadF2/MadR2) permitió obtener un nuevo ácido nucleico de 604 pb.

La secuencia nucleotídica del cebador MadF2 es:

5'-TTCACCTTGTTGGAACCGCTGGG-3'

La secuencia nucleotídica del cebador MadR2 es:

5'-TCATTACCGAGCCGACCTGC-3'

Las sondas se marcan con dUTP-digoxigenina (*DIG DNA Labelling and Detection Kit*, Roche Diagnostics, Laval Canadá).

EJEMPLO 3: CLONACIÓN DE UN ÁCIDO NUCLEICO DE LA SEQ ID NO:11

3.1. Transferencia Southern

Los experimentos de transferencia Southern se realizaron en el DNA genómico de la cepa *M. austroafricanum* I-2562 tal como se preparó en el EJEMPLO 1. 5 µg de DNA genómico se digirió parcialmente con ayuda de enzimas de restricción elegidas en función de su sitio de restricción *Bam*HI, *Eco*RI, *Fsp*I, *Kpn*I, *Nhe*I, *Nru*, *Pst*I, *Pvu* II o *Sma*I (Pistas 2-8, respectivamente en la Figura 5). Después de migración en gel, los DNA digeridos se transfirieron en una membrana de nilón. Con ayuda de la sonda de 204 pb marcada con dUTP-digoxigenina (*DIG DNA Labelling and Detection Kit*, Roche Diagnostics, Laval, Canadá), se procedió al cribado por hibridación de los ácidos nucleicos digeridos. La sonda de 204 pb se hibridó específicamente con fragmentos de ácidos nucleicos comprendidos entre 2 y 12 kb (Figura 5). Las posiciones correspondientes a la migración del DNA del marcador de peso molecular III marcado con digoxigenina (Roche; Pista 1 y Pista 9) están indicadas en la parte izquierda de la figura).

3.2. Construcción de mini-bancos, cribado y clonación

Estos fragmentos de ácidos nucleicos se clonan en el plásmido pBluescript II KS (pBKS) (+/-) después de digestión por *Sma*I. Los plásmidos recombinantes se utilizan para transformar bacterias competentes DH10B de *E.coli* preparadas (Hanahan, D et al., *Methods Enzymol.* 1991, **204**:63-113). Cada colonia bacteriana procedente de un clon bacteriano que contiene un plásmido recombinante se transfiere a un membrana de nitrocelulosa y se criba por hibridación en colonia con la sonda de 604 pb, tal como se obtiene en el EJEMPLO 2.3. Para cada clon positivo, es decir que se hibrida con la sonda, el plásmido recombinante (pKS1) se extrajo y purificó gracias al kit QIAprep Spin Miniprep (QIAGEN, Mississauga, Canadá). Debido a la fuerte inestabilidad del fragmento *Sma*I en el seno del plásmido pBKS, que implica la aparición de plásmidos de tamaños variados, se realizó una digestión por *Sma*I de 10 µg de plásmido pBKS 1 seguida por una extracción en gel. La extracción de la banda *Sma*I se obtuvo gracias a una extracción después de migración en un gel «de bajo punto de fusión». Se obtuvieron aproximadamente 600 a 1000 pares de bases por cada extracción, por consiguiente fueron necesarias una docena de extracciones para obtener una secuencia completa. Después del análisis de la secuenciación, faltaba la parte hacia el extremo 5' del gen *mpdC*, se realizó un nuevo minibanco con el fin de obtener un fragmento *Pst*I de 5,6 kb utilizando las mismas técnicas que precedentemente, pero utilizando la sonda de 914 pb que se encuentra en el primer inserto y obtenida con ayuda del par de cebadores MF3/MR3. El plásmido pKS3 se aisló a partir de un clon positivo revelado por «hibridación en colonias» (Figura 6). El fragmento *Pst*I-*Pst*I detectado por la sonda fue pues clonado según el mismo protocolo que precedentemente para obtener un plásmido recombinante pKS3.

La totalidad del inserto *Pst*I presente en este plásmido se secuenció según la técnica de «primer-walking (recorrido del cebador)» con ayuda de una T7 DNA-polimerasa (*T7-DNA sequencing kit*, Applied Biosystems, Foster City, CA) y de un secuenciador (secuenciador con fluorescencia automático *ABT prism 377*, Applied Biosystems, Foster City, CA). Las secuencias nucleotídicas y de aminoácidos correspondientes se compararon con las contenidas en las bases de datos EMBL, *Swissprot* y *GenBank* utilizando los programas informáticos BLASTN y BLASTX del *National Center for Biotechnology Information (NCBI)*.

EJEMPLO 4: IDENTIFICACIÓN Y CARACTERIZACIÓN DE LOS DIFERENTES MARCOS DE LECTURA ABIERTOS (ORF) COMPRENDIDOS EN LA SECUENCIA SEQ ID NO:11 DE M. AUSTRORAFRICANUM I-2562

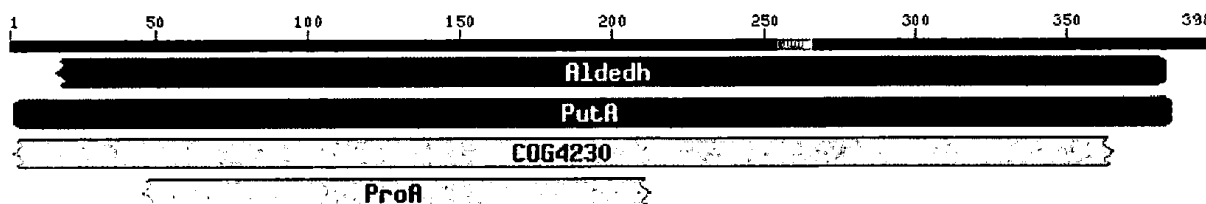
5 El análisis de los marcos de lecturas abiertos presentes en la secuencia de DNA *Pst*I/*Sma*I, SEQ ID NO:11 obtenida después de las dos clonaciones descritas en el EJEMPLO 3 se realizó con ayuda del programa informático «*ORF finder*». Se identificaron cinco secuencias codificantes SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7 y SEQ ID NO:9; dichas secuencias codifican respectivamente los polipéptidos SEQ ID NO: 2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8 y SEQ ID NO:10 (véase la tabla siguiente).

| Secuencia nucleotídica | ORF | Secuencia de aminoácidos del polipéptido | Nombre del polipéptido | Peso molecular |
|------------------------|-------------|--|------------------------|----------------|
| SEQ ID NO:1 1515 pb | <i>mpdC</i> | SEQ ID NO:2 552 aminoácidos | MpdC | 55 kDa |
| SEQ ID NO:3 648 pb | <i>orf1</i> | SEQ ID NO:4 215 aminoácidos | | 27 kDa |
| SEQ ID NO:5 1659 pb | <i>mpdB</i> | SEQ ID NO:6 552 aminoácidos | MpdB | 64 kDa |
| SEQ ID NO:7 672 pb | <i>orf2</i> | SEQ ID NO:8 223 aminoácidos | | 25 kDa |
| SEQ ID NO:9 1233 pb | <i>mpdR</i> | SEQ ID NO:10 410 aminoácidos | MpdR | 47 kDa |

10 La secuencia de aminoácidos SEQ ID NO:6 del polipéptido MpdB de 64 kDa por tanto pudo ser correlacionada con las primeras secuencias polipeptídicas y las amplificaciones por PCR de 204 y 604 pb. Este gen *mpdB* está comprendido en una organización genética que comprende otros 3 ORF en la misma orientación, *mpdC*, *orf1* y *orf2*, que codifican respectivamente las proteínas de 55, 27 y 25 kDa. Aguas arriba (o sea en dirección 5') de este grupo, se identificó un nuevo gen orientado en el sentido opuesto al primer grupo y que codifica una proteína de 47 kDa y se denominó *mpdR*. La organización de este sistema se esquematiza en la Figura 7.

15 4.1. ORF *mpdC*, SEQ ID NO:1

El análisis por el programa BLAST esquematizado a continuación muestra fuertes similitudes con numerosas deshidrogenasas que pertenecen a la familia de los aldehído-deshidrogenasas. El polipéptido SEQ ID NO:2 codificado por este ORF corresponde a un polipéptido de 55 kDa, cuya expresión es inducida en MTBE. Este polipéptido es capaz de deshidrogenar el hidroxibutiraldehído a HIBA.



20 Resultado obtenido por programa informático "NCBI domain search":

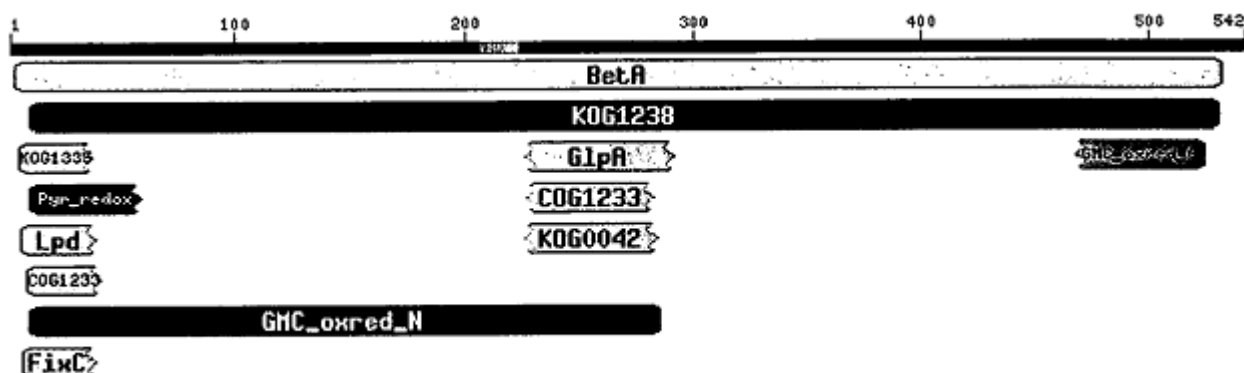
| | |
|---|------------------|
| Aldehdh , Familia de las aldehído-deshidrogenasas | e-valor = 3e-114 |
| PutA , aldehído-deshidrogenasas dependientes de NAD | e valor = 5e-106 |
| COG4230 , Delta-1-pirrolina-5-carboxilato-deshidrogenasa | e valor = 3e-44 |

4.2. ORF *orf1*, SEQ ID NO:3

El polipéptido SEQ ID NO:4 codificado por este ORF corresponde a un polipéptido de 27 kDa, cuya expresión es inducida en MTBE.

4.3. ORF *mpdB*, SEQ ID NO:5

5 El polipéptido SEQ ID NO:6 codificado por este ORF corresponde a un polipéptido de 64 kDa, cuya expresión es inducida en MTBE. Hay fuertes similitudes con numerosas oxidasas y deshidrogenasas que pertenecen a la familia de las GMC (Glucosa-Metanol-Colina)-óxido-reductasas. El análisis por el programa "NCBI Conserved Domain Search" detectó más precisamente similitudes con la proteína codificada por el gen *betA*, que codifica la colina-deshidrogenasa que pertenece a esta misma familia. La secuencia de aminoácidos deducida puede ser correlacionada con la secuencia N-terminal y las dos secuencias internas de la proteína de 64 kDa inducida específicamente en MTBE y responsable de la transformación del 2-M-1,2-PD en hidroxibutiraldehído.



Resultado del programa informático NCBI conserved domain search:

Resultado obtenido para una búsqueda de dominios conservados por NCBI

15

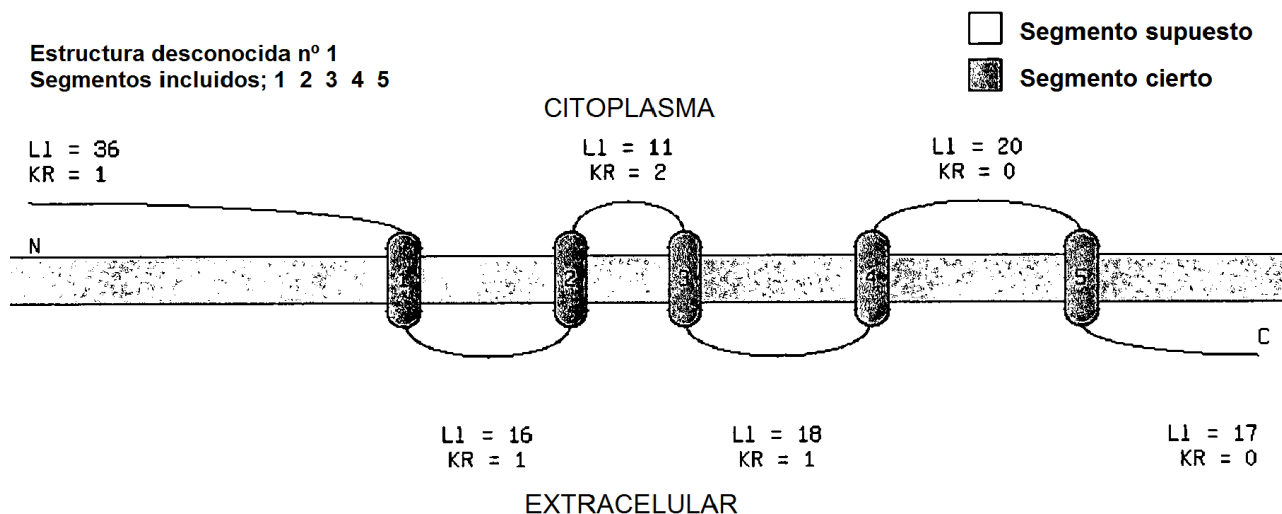
| | |
|---|-----------------|
| BetA , Colina-deshidrogenasa y flavoproteínas relacionadas | e-valor = 3e-48 |
| KOG1238 , Glucosa-deshidrogenasa/colina-deshidrogenasa | e-value = 2e-07 |
| KOG1335 , Dihidrolipoamida-deshidrogenasa [Producción de energía]... | e-value = 6e-05 |
| GMC_oxred_N , GMC óxido-reductasa | e-value = 2e-03 |
| GMC_oxred_C , GMC óxido-reductasa | e-value = 4e-03 |

4.4. ORF *orf2*, SEQ ID NO:7

20 El polipéptido SEQ ID NO:8 codificado por este ORF corresponde a un polipéptido de 25 kDa. El análisis por el programa BLAST reveló una homología con una di/tripéptido permeasa (COG3104: Dipéptido/tripéptido-permeasa [*Rubrivivax gelatinosus* PM1] Longitud = 372, Valor esperado = 2e-24). El análisis de la hidropatía de esta proteína por el programa "Predict Protein" reveló la presencia de 5 segmentos transmembranales cuya organización se esquematiza a continuación:

La expresión de este polipéptido es inducida probablemente en presencia de MTBE.

Esta proteína se encuentra pues en el seno de la membrana de las bacterias de la especie *M. austroafricanum*.



4.5. ORF *mpdR*. SEQ ID NO:9

El polipéptido SEQ ID NO:10 codificado por este ORF corresponde a un polipéptido de 47 kDa. Este ORF está situado en dirección 5' respecto a las cuatro ORF precedentes y está orientado en el otro sentido como muestra la Figura 7. El análisis por el programa BLAST reveló una similitud del primer resto de la secuencia con numerosos reguladores transcripcionales $\sigma 54$, principalmente del tipo AcoR:

| | |
|--|-----------|
| Activador transcripcional dependiente de Sigma-54 [<i>Azoarcus</i> sp. EbN1] | e = 6e-23 |
| COG3284: Activador transcripcional AcoR del metabolismo de acetoina/glicerol | e = 1e-22 |

Después de estos análisis no es posible detectar si la expresión de este polipéptido es inducida específicamente o no en el perfil MTBE, pero es posible detectar su expresión en el momento del análisis por RT-PCR. Dicha expresión desempeña potencialmente un papel en la regulación de la expresión de los genes *mpd*.

EJEMPLO 5: EXPRESIÓN POR RT-PCR DEL GRUPO DE LOS GENES *mpd*

5.1. Extracción del mRNA de *M. austroafricanum* I-2562

Para los experimentos de extracción de RNA, todas las herramientas, todas las soluciones y el recipiente se prepararon de manera que se evitara la contaminación por RNAasas según los métodos estándares (Sambrook et al., *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, Nueva York, 2001)

El RNA total de *M. austroafricanum* I-2562 se extrajo a partir de 20 mL de un cultivo a D.O._{600nm} de 0,7. Después de una incubación en hielo de 30 minutos, las células se centrifugaron (8000 x g durante 10 minutos a 4°C) y se conservaron sobre hielo. A continuación, se añadieron al sedimento de células 600 μ L de lisozima (Roche Diagnostics, Laval, Canadá) de 3 mg.mL⁻¹ y 600 μ L de lisostafina (Sigma, St. Louis, Missouri) de 50 μ g.mL⁻¹ y la mezcla se incubó a 37°C durante 10 minutos. A continuación se añadieron 10 mL de RNAwiz™ (Ambion, Austin, Texas) y la mezcla se sometió a agitación con vórtice durante 15 segundos. A continuación la mezcla se fraccionó en tubos de 2 mL (14000 μ L/tubo) que contenían 250 mg de bolas de zirconio-sílice (0,1 mm de diámetro) colocadas en un adaptador para vórtice (Ambion) y se sometieron a agitación con vórtice durante 10 minutos. Después de haber centrifugado (1300 x g a 4°C durante 5 minutos), el lisado bacteriano correspondiente a los líquidos sobrenadantes de cada tubo se transfirió a nuevos tubos de 2 mL y se añadieron 2 mL y 0,2 volúmenes de CHCl₃ en cada tubo. Los tubos se sometieron a continuación a agitación con vórtice durante 30 segundos, se incubaron a temperatura ambiente durante 10 minutos, y después se centrifugaron (13000 x g a 4°C durante 5 minutos) con el fin de recuperar los líquidos sobrenadantes en los nuevos tubos de 2 mL. Los extractos de RNA se precipitaron añadiendo secuencialmente 0,5 volúmenes de H₂O tratada con DEPC, 1 volumen de isopropanol y 1/50 de volumen de glicógeno exento de RNasa de 5 mg.mL⁻¹ (Ambion), se mezclaron bien y se incubaron a temperatura ambiente durante 10 minutos. El RNA se centrifugó (13000 x g a 4°C durante 5 minutos), se lavó dos veces con etanol del 70%, se secó 30 segundos (DNA Speed Vac, Savant) y se volvió a suspender en H₂O-DEPC (25 mL/tubo). La calidad del RNA se determinó por electroforesis con ayuda de un gel de agarosa al 1% no desnaturalizante en un tampón TBE 0,5X (Trizma Base 45 mM, ácido bórico 45 mM, Na₂EDTA 1 mM a pH=8,3). El RNA se precipitó con 1/10 volumen de acetato de sodio 3M de pH=7,0 y 2,5 volúmenes de etanol al 95%, y después se conservó a 80°C.

5.2. Tratamiento con DNasa y transcriptasa inversa - PCR (RT-PCR)

5 Dos mg de extractos de RNA de *M. austroafricanum* I-2562 se trataron con el kit «DNA-Free» (Ambion, Austin, Texas) según las recomendaciones del fabricante. La eliminación del DNA se verificó con ayuda de amplificación por PCR utilizando los cebadores específicos creados en el grupo *mpd* y por electroforesis en gel de agarosa al 1%. La concentración final del RNA se cuantificó a 260 nm con ayuda de un Nanodrop (Nanodrop Technologies, Wilmington, USA). Los experimentos de transcriptasa inversa-PCR (RT-PCR) se emplearon según el protocolo del kit «One step RT-PCR» de QIAGEN (QIAGEN, Mississauga, Canadá) tal como se describe por el fabricante. Se utilizaron 120 ng de RNA en cada tentativa de RT-PCR con, cada vez en paralelo, una amplificación por PCR con control positivo (y negativo) con (o sin) 100 ng de gDNA de *M. austroafricanum* I-2562. Las condiciones de RT-PCR se describen a continuación.

- etapa de transcripción inversa:

1) 30 minutos a 50°C

2) ∞ a 50°C

- etapa de PCR:

15 15 minutos a 95°C

30 a 40 ciclos (variables en función de la intensidad de la señal):

1) 30 segundos a 94°C

2) 1 minuto a 55°C

3) 1 minuto a 72°C

20 - etapa de elongación final:

1) 10 minutos a 72°C

2) ∞ a 4°C

5.3. Expresión por RT-PCR del grupo de los genes *mpd* durante el crecimiento de *M.austroafricanum* I-2562 en TBA

25 La expresión de los genes *mpd* es analizada después del crecimiento de las bacterias en TBA, el principal intermedio de la biodegradación del MTBE (véase la Figura 1). Se crean pares de cebadores que se encuentran en los diferentes genes *mpd* (véase la Tabla) siguiente con el fin de verificar que el mRNA que corresponde a este grupo era transcrito en un monocistrón. El RNA total se extrae de los cultivos después del crecimiento en TBA y las RT-PCR se efectúan con los diferentes cebadores. El resultado de esta experimentación que se presenta en la Figura 8 muestra que muy probablemente la transcripción del grupo de los genes *mpd* es monocistrónica; estando los genes organizados en operón. (Ma et al. *J. Bacteriol.*, 2002, **184**: 5733-5745).

30

Tabla: Cebadores utilizados en el conjunto de experimentos de amplificación por PCR y RT-PCR:

| Cebadores | Secuencia | Región amplificada |
|--------------------------|--------------------------------|---------------------|
| Bott1 (SEQ ID N :15) | 5'-TGCACACAGGCCACAACCCA-3' | rDNA 16S |
| Bott2 (SEQ ID N :16) | 5'-GAGAGTTTGATCCTGGCTCAG-3' | rDNA 16S |
| 244 (SEQ ID N :17) | 5'-CCCCTGCTGCCTCCCGTAG-3' | rDNA 16S |
| Tb11 (SEQ ID N :18) | 5'-ACCAACGATGGTGTGTCCAT-3' | <i>hsp65</i> |
| Tb12 (SEQ ID N :19) | 5'-CTTGTGCAACCGCATAACCCT-3' | <i>hsp65</i> |
| MaFV2 (SEQ ID N :20) | 5'-GTCTAATACCGAATACACCCTTCT-3' | rDNA 16S |
| MaRV6 (SEQ ID N :21) | 5'-GTAGTTGGCCGGTCCTTCTTCTCC-3' | rDNA 16S |
| MF1 (SEQ ID N :22) | 5'-TGAGAAGCCTCGTGTATTAC-3' | <i>orf1-mpdB</i> |
| MR1 (SEQ ID N :23) | 5'-GAGATAAGGCGTGGTGAA-3' | <i>orf1-mpdB</i> |
| MF2 (SEQ ID N :24) | 5'-AGTGACGGCACCCATAAGTG-3' | <i>mpdR</i> interna |
| MR2 (SEQ ID N :25) | 5'-TCGAGGTGTTGAGGTCCGAAT-3' | <i>mpdR</i> interna |
| MF3 (SEQ ID N :26) | 5'-ATCATCCCGTGGAACTAC-3' | <i>mpdC</i> interna |
| MR3 (SEQ ID N :27) | 5'-TGACCTGGGCGATGTGTT-3' | <i>mpdC</i> interna |
| MF4 (SEQ ID N :28) | 5'-ATCAGACCTGGGATGTGC-3' | <i>mpdB-orf2</i> |
| MR4 (SEQ ID N :29) | 5'-GGCTGTGAAAGTCGGATGA-3' | <i>mpdB-orf2</i> |
| RT-PCR-F1 (SEQ ID N :30) | 5'-AGTGACGGCACCCATAAGTG-3' | <i>mpdR</i> interna |
| RT-PCR-R1 (SEQ ID N :31) | 5'-TCGAGGTGTTGAGGTCCGAAT-3' | <i>mpdR</i> interna |
| RT-PCR-F2 (SEQ ID N :32) | 5'-GCAGGTCGGCTCGGTAATGA-3' | <i>mpdC-orf1</i> |
| RT-PCR-R2 (SEQ ID N :33) | 5'-GTAATACACGAGGCTTCTCA-3' | <i>mpdC-orf1</i> |
| RT-PCR-F3 (SEQ ID N :34) | 5'-ACGGTCTCGTCCGCAAATAC-3' | <i>mpdB-orf2</i> |
| RT-PCR-R3 (SEQ ID N :35) | 5'-GCACATCCAGGTCTGAT-3' | <i>mpdB-orf2</i> |

6.1. Construcción del vector de expresión que contiene los genes *mpd* y transformación de las bacterias

- 5 Los genes comprendidos en la organización genética *mpd* fueron insertados en el seno del vector pCL4D en dos etapas (Picardeau et al., *Microbiology*, 2000, **146**: 305-313). En un primer tiempo el fragmento *NotI-NotI* de 4401 pares de bases contenido en el plásmido pKS 1 se digirió por *NotI* (NEB, Pickering, Canadá) y se clonó en el sitio *HindIII* de pCL4D. Las células quimiocompetentes de *E. coli* DH10B se utilizaron para la transformación y la selección de los transformantes en el medio LB que contiene 20 $\mu\text{g}\cdot\text{mL}^{-1}$ de kanamicina. La detección de los clones positivos que contenían el plásmido recombinante p4D1 se realizó por la amplificación por PCR con ayuda de los cebadores *Forward* y *Reverse* (Tabla en el EJEMPLO 5). En la segunda etapa el fragmento *PstI* de 5574 pb que contenía la parte que falta del gen *mpdC*, el gen *mpdR* y el *orf3* se extrajo por digestión del plásmido pKS3 con ayuda de la enzima *PstI*. Este fragmento se introdujo en el seno del plásmido p4D1 digerido previamente por *PstI* que permite así llevarse el fragmento *PstI* correspondiente a la sonda de 914 pb presente en el plásmido p4D1. La orientación del fragmento *PstI* en el plásmido recombinante p4D2 que contiene la totalidad del grupo *mpd* se verificó por PCR.
- 10
- 15 Con el fin de expresar el conjunto de los genes clonados, el vector p4D2 se introdujo en células competentes de la cepa *M. smegmatis* mc2 155 utilizando la técnica de electroporación. Un transformante (cepa *M. smegmatis* mc2 155-clon 9) procedente de esta transformación se aisló en una caja con LB que contenía kanamicina (20 $\mu\text{g}\cdot\text{mL}^{-1}$). El plásmido pCL4D que no comprende inserto fue introducido igualmente en bacterias de la cepa *M. smegmatis* mc2 155 (cepa *M. smegmatis* mc2 155-pCL4D) y sirvió de control.

20 6.2. Expresión funcional de los genes *mpdB* y *mpdC* en *M. smegmatis* mc2 155

- Las bacterias *M. smegmatis* mc2 155-clon 9 (con inserto) y *M. smegmatis* mc2 155-pCL4D (control) se conservan en el medio LB que contienen 20 $\text{mg}\cdot\text{L}^{-1}$ de kanamicina. Las dos cepas se cultivan en 200 mL de medio LB que contiene 20 $\text{mg}\cdot\text{L}^{-1}$ de kanamicina durante 72 horas a 30°C. Las células de *M. smegmatis* mc2 155-clon 9 y *M. smegmatis* mc2 155-pCL4D se recogen por centrifugación (13.000 g durante 15 minutos), se lavan dos veces y se ponen en suspensión en 40 mL de tampón de fosfato (20 mM, pH 7) que contiene el sustrato a ensayar (MTBE, TBA, 2-M-1,2-
- 25

PD o HIBA) en frascos sellados de 120 mL. Después de la siembra los frascos se incuban a 37°C en un agitador rotatorio. Cuando es necesario, se añade cloranfenicol a partir de una solución en agua esterilizada por filtración (0,22 µm) con el fin de obtener una concentración final de 400 mg.L⁻¹. Las muestras filtradas de estos cultivos son analizadas por CPG o HPLC. La degradación del sustrato es seguida por un periodo de 24 horas. Las actividades específicas (mg de sustrato degradado. g⁻¹ de biomasa.h⁻¹) se calculan a partir de las velocidades máximas de degradación.

6.3. Expresión en *M. smegmatis* mc2 155 del grupo de los genes *mpd* aislados de *M. austroafricanum* I-2562

Un fragmento de 9,1 kb, SEQ ID NO:11 que contiene el grupo de los genes *mpd* de *M. austroafricanum* I-2562 se clonó en el plásmido pCL4D (Picardeau et al., 2000). Este plásmido, denominado p4D2, se utilizó seguidamente para efectuar la transformación de *M. smegmatis* mc2 155. La transformación se realizó con el plásmido pCL4D, como control. Se seleccionaron dos transformantes: *M. smegmatis* mc2 155-clon 9 (que contenía p4D2 y que lleva por tanto los genes *mpd*) y *M. smegmatis* mc2 155-pCL4D (que lleva el vector, pCL4D). Las dos cepas se cultivan en el medio LB que contiene kanamicina y después se ensayan en dos experimentos en células en reposo para analizar la capacidad de degradar el MTBE, el TBA, el 2-M-1,2-PD y el HIBA. No se observó ninguna degradación del MTBE, del TBA o del HIBA con las cepas *M. smegmatis* mc2 155-clon 9 y *M. smegmatis* mc2 155-pCL4D (Tabla 2).

La degradación del 2-M-1,2-PD y la producción estequiométrica de HIBA (312,1 ± 2,3 µM degradado y 311,6 ± 4,5 µM producido, respectivamente) se observó solamente en presencia de la cepa *M. smegmatis* mc2 155-clon 9 (Figura 9). No se observó ninguna degradación del 2-M-1,2-PD en presencia de la cepa *M. smegmatis* mc2 155-pCL4D o en el control abiótico. Al ser las dos cepas cultivadas en medio completo de LB, la inducción de los genes era necesaria para producir las enzimas correspondientes y esto es porque la degradación no comienza más que después de 4 horas de incubación en presencia de 2-M-1,2-PD. No se observa ninguna degradación del 2-M-1,2-PD cuando las células de *M. smegmatis* mc2 155-clon 9 se incuban en presencia de 2-M-1,2-PD y de cloranfenicol que se sabe que puede inhibir la traducción de los mRNA a proteínas.

Se calculó la velocidad máxima de degradación del 2-M-1,2-PD por la cepa *M. smegmatis* mc2 155-clon 9 y es 2,34 ± 0,41 µmol.g⁻¹ (peso seco).min⁻¹.

EJEMPLO 7: IDENTIFICACIÓN DE UNA NUEVA CEPA DE *M. AUSTROAFRICANUM* I-3401, CAPAZ DE CRECER EN MTBE

El agua de lluvia recogida en el fondo de un depósito de almacenamiento de una gasolina aditivada con MTBE se utilizó para sembrar el medio mineral MM que contenía 200 mg.L⁻¹ de MTBE. El MTBE fue lentamente pero totalmente consumido en 140 días. Todas las bacterias que forman colonias después de esparcimiento en cajas de medio LB, se re-aislan individualmente en placas y después se analizan para ver su capacidad de crecer en MM que contiene MTBE como única fuente de carbono y de energía. Se aisló una bacteria Gram-positiva y aerobia estricta que forma bastoncillos, y que crece en forma de colonias amarillas en cajas de medio LB y se reveló capaz de de crecimiento en MTBE.

El rDNA 16S de la cepa I-3401 fue totalmente secuenciado para las dos cadenas y el gen *hsp65* fue parcialmente secuenciado. La secuencia del rDNA 16S muestra que la nueva cepa I-3401, SEQ ID NO:13, es muy próxima a *M. austroafricanum* I-2562, con siete nucleótidos diferentes (5 sustituciones y 2 inserciones en I-3401). La secuencia del gen *hsp65* difiere en ocho nucleótidos de la del gen *hsp65* de la cepa-tipo (de referencia) de *M. austroafricanum*. Se ha mostrado la capacidad de *M. austroafricanum* I-3401 de crecer en MTBE (Figura 10).

EJEMPLO 8: CREACIÓN DE CEBADORES ESPECÍFICOS DEL rDNA 16S DE *M. AUSTROAFRICANUM* Y CONDICIONES ESPECÍFICAS DE LA PCR

8.1. Análisis del rDNA 16S y del gen *hsp65* de la cepa *M. austroafricanum* I-2562

Con el fin de identificar la nueva cepa I-3401, se realizó una amplificación por PCR del rDNA 16S utilizando la pareja de cebadores Bott 1 *forward* / Bott 2 *reverse* (véase tabla del EJEMPLO 5). El producto de amplificación se purificó utilizando el kit Qiagen (Qiagen, Mississauga, Ontario, Canadá) y se realizó la secuenciación nucleotídica utilizando el cebador 244 (véase tabla del EJEMPLO 5). La amplificación por PCR del gen *hsp65* se realizó utilizando el cebador *forward* Tb 11 y el cebador *reverse* Tb12 (véase tabla del EJEMPLO 5). El producto de amplificación se purificó con el kit Qiagen y la secuenciación se efectuó utilizando el cebador Tb 11. La reacción de secuenciación se efectuó con el kit BigDye *ciclo terminator* (Version 3.1, Applied Biosystems, Foster City, CA, USA) como se describe por el fabricante utilizando 25 ng de DNA purificado y 15 µmol de los cebadores universales utilizados para la secuenciación del rDNA 16S de las Eubacterias. La reacción se programa como sigue: 25 ciclos de 10 segundos a 96°C, 5 segundos a 50°C, y 4 minutos a 60°C. Los productos de la secuenciación se purifican en columnas Centri-Sep como se ha descrito por el fabricante (Princeton Separations, Inc., Adelphia, NJ, USA) para eliminar el exceso de terminadores. Las secuenciaciones se realizan en un secuenciador automático con detección por fluorescencia ABT Prism 377 (Applied Biosystems, Foster City, CA). Las secuencias nucleotídicas obtenidas se comparan con las contenidas en el banco de datos de genes EMBL/GenBank y las identidades se evalúan utilizando el sistema de alineamiento BLAST (Altschul et al., 1997).

8.2. Creación de cebadores y condiciones específicas de PCR

Se alinean las secuencias del rDNA 16S de *M. austroafricanum* I-2562 y de especies próximas de micobacterias y se comparan las secuencias conservadas y variables. Sobre la base de los análisis de estos alineamientos múltiples, Se comparan varios pares de cebadores específicos de la especie *M. austroafricanum* con las secuencias del rDNA 16S disponibles utilizando el programa de búsqueda en los bancos de datos BLAST (Altschul et al., *Nucleic Acids Res.*, 1997, **25**: 3389-3402). Estas secuencias se analizaron también con el fin de determinar su temperatura de desnaturalización (T_m), la posibilidad de formación de dímeros y su contenido en (G+C) utilizando el programa Amplify. La pareja de cebadores más eficaz era: MaFV2 *forward* y MaRV6 *reverse* (véase Tabla del EJEMPLO 5). El tamaño del producto de PCR esperado es 331 pb.

Las PCR se realizan de la manera siguiente: cada tubo contiene 5 μ L de DNA genómico que procede de una lisis por calentamiento ("*boiling lysis*") durante 10 minutos de colonias aisladas en cajas recientemente esparcidas con las cepas de colección, las cepas de *Mycobacterium austroafricanum* o las muestras procedentes de una biobarrera sembrada con *M. austroafricanum* I-2562, 2,5 unidades de DNA-polimerasa *Taq* (Amersham Pharmacia Biotech), 5 μ L del tampón de DNA-polimerasa *Taq* diluido 10X (Amersham Pharmacia Biotech), 25 pmol de cada uno de los cebadores, 4 μ L de desoxirribonucleótidos-trifosfatos a 2,5 mM (a 200 pM cada uno: dATP, dGTP, dCTP y dTTP), 2 μ L de $MgCl_2$ a 25 mM, y de agua destilada estéril para tener un volumen final de 50 μ L. El control negativo contiene la misma mezcla que se ha descrito antes salvo que el DNA es reemplazado por agua estéril. Las muestras se calientan previamente 3 minutos a 95°C en un aparato Bio-Rad Thermal iCycler (Bio-Rad, Mississauga, Ontario, Canadá) y después la temperatura se baja a 80°C antes de su adición a la mezcla enzimática DNA-polimerasa *Taq*/tampón diluido 10X. Con el fin de obtener condiciones de amplificación específica, las condiciones de amplificación son las siguientes: 30 segundos a 94°C, 1 minuto a 68°C, 1 minuto a 72°C repetido durante 30 ciclos y la extensión final se efectúa a 72°C durante 7 minutos. Los tubos de PCR se conservan a 4°C hasta el momento de efectuar la electroforesis en gel de agarosa. 10 μ L de cada uno de los productos de la PCR se mezclan con 2 μ L de tampón de deposición que contiene glicerol al 30% (v/v), azul de bromofenol al 0,15% (p/v), xileno-cianol al 0,15% (p/v). Los diferentes productos de la PCR así preparados y el marcador de DNA de 1 kb *GeneRuler*TM (MBI Fermentas, Inc., Burlington, Ontario, Canadá) se depositan en un gel de agarosa al 1% (p/v) preparado en TAE. Después de la migración, el gel se colorea con bromuro de etidio y se revela con rayos UV a 254 nm sobre un transiluminador. Los geles se fotografían utilizando un película Polaroid de tipo 57.

EJEMPLO 9: DETECCIÓN DE LA PRESENCIA DE LA ORGANIZACIÓN GENÉTICA *mpd* EN EL SENO DE LA NUEVA CEPA *M. AUSTROAFRICANUM* I-3401.

Se ha estudiado la biodegradación del MTBE por la cepa *M. austroafricanum* I-3401 al ser similar a la de *M. austroafricanum* I-2562, la presencia o no del «grupo» *mpd* en el seno de la nueva cepa. Así se crearon parejas de cebadores, MF1/MR1, MF2/MR2, MF3/MR3 y MF4/MR4 que amplifican específicamente en los diferentes genes del grupo *mpd* (Tabla del EJEMPLO 5). Se realizaron amplificaciones por PCR efectuadas sobre el DNA genómico de *M. austroafricanum* I-3401 o I-2562 (este último como control positivo) (Figura 11) y muestran amplificaciones positivas para todas las parejas de cebadores. Esto demuestra la presencia de una organización genética similar a la de los genes *mpd* de *M. austroafricanum* I-2562 en el seno del genoma de la nueva cepa *M. austroafricanum* I-3401. Es pues cierto que la nueva cepa utiliza el mismo camino de reacción para asimilar el 2-M-1,2-PD en el curso del catabolismo del MTBE.

EJEMPLO 10: DETECCIÓN ESPECÍFICA DE *M. AUSTROAFRICANUM* I-2562 Y I-3401 POR PCR.

Se crearon los cebadores MaFV2 y MaRV6 específicos de la especie *M. austroafricanum* (véase la tabla del EJEMPLO 5). Se evaluó la especificidad de estos cebadores fijándose en dos regiones variables V2 y V6 del rDNA 16S de *M. austroafricanum* I-2562 utilizando el DNA genómico de cepas más o menos próximas. Los resultados presentados en la Figura 12 muestran que solo los DNA genómicos de dos cepas de *M. austroafricanum*, I-2562 y I-3401, que metabolizan el MTBE permiten obtener una amplificación por PCR positiva. Estos cebadores permiten pues detectar específicamente los microorganismos de esta especie. Otros microorganismos del género *Nocardiaceae* no presentan amplificación por PCR.

EJEMPLO 11: COMPARACIÓN ENTRE LAS CEPAS I-2562 y I-3401

| | <i>M. austroafricanum</i> | |
|--|--|---|
| | I-2562 | I-3401 |
| Capacidad de degradación del ETBE (130 mg/L) | 100 días | 33 días |
| Metabolismo del MTBE: asimilación de los intermedios | . | acumulación del 2-M-1,2-PD |
| Velocidad de degradación del 2-metil-1,2 propanodiol | 413 mg.g ⁻¹ . peso seco.h ⁻¹ | 60 mg.g ⁻¹ peso seco.h ⁻¹ |

5 Una diferencia importante entre la cepa I-2562 y la cepa I-3401 reside en su capacidad de degradación del ETBE: para una concentración equivalente de ETBE (130 mg/L), *M. austroafricanum* I-2562 degrada el ETBE en 100 días mientras que I-3401 realiza esta degradación tres veces más rápido. Por otra parte, la velocidad de degradación del 2-M-1,2-PD cuando se utiliza directamente como sustrato es aproximadamente 7 veces más rápida en I-2562 que en I-3401. Finalmente, es posible detectar la presencia de 2-M-1,2-PD en el curso de la degradación del alcohol *terc*.butílico (TBA) en I-3401 debido al hecho de esta diferencia de velocidad de degradación del 2-M-1,2-PD.

LISTA DE SECUENCIAS

<110> Institut Français du Pétrole National Research Council of Canada

5 <120> Polipéptidos que tienen una actividad en la vía de degradación del metil-*terc*.butil-éter (MTBE) y sus utilizations

<130> BCT060237

10 <160> 35

<170> PatentIn versión 3.3

<210> 1

15 <211> 1515

<212> DNA

<213> Mycobacterium austroafricanum

<400> 1

```

atgacccgaa ctctctcggc tgatgccgac acccgcacgg cgacacctcc gctgatgtac      60
gtcaacggcg agtggctgcc cgcccgcagt ggggccacct ttcccacat cgaacccagc      120
acgggtcgcgac cgatcactga gattccccgc ggggactcga gcgatgtgga cgcggcggtg      180
aaagccgctg ccgacgtggc cgttgagtgg cagttcaccg atgccatcac ccgcgccgcc      240
ctgctcaggc gattggcggg gctggtgcca gagaacgccg aggagtggc gcggatcgag      300
tcgctggact cgggtcacta tctggcgaag gcgcgtgaac tggtgaccgc gatacccctg      360
tggctcgagt actgggccgg cgacgtgac aaagtgggcg gccgcacat cgctgtaccg      420
ggtaacaaac tcagcttcac cttgttgaa ccgctgggcg tcaccgcga catcatcccg      480
tggaactacc cgctgttgat ccttgctcgg tccatcgccc cggcactcgc attgggcaac      540
acctgtgtcg tcaagccgc tgaggacacg tccctgtcag cgctgaagt cgccgagctg      600
gtacacgccg ccggttttcc cgccggagtg ttcaatgtgg tgaccggtta cggttccgaa      660
gccggcgcgg ctcttgccgc tcaccccgag gtgcgcggaa tcaccttcac cggttcgacc      720
gagaccgggc gggagatcgc ccgactgggc ggccaacaca tcgccaggt caacctggaa      780
ctcggcggga agagcccgtt ggtcgtcttt cccgacgcgc cgctcgaaga cgccgtagag      840
gtggccgtac agggcttctg ctcacgggca gggcaagtgt gtgtcgccgg gagccggctc      900
ttctccatg aggacatcgc cgaccggttc ctcgagatgc tcgtttcgcg actcgagact      960
gtcaccgtcg gacacccgtt cgacgggtgc acccagatgg gtccgctcgc ctcgaagaag     1020
cactacgacc gtgtgcgtga gtacatcgag gtcgggaagc aggaggcgac cctgctctac     1080
ggcggcggtc ggccgacgga cacgccgat gacgggttct tcgtcgagcc aacggttttc     1140
gtcgacgtcg caacggatgc gcggatcgca cgcgaggaga tcttcgggcc cgtcacagcg     1200
gtgatgcggt ggtcatcggc cgacgatctg atcgccacca tcaatgatc ggaattcggc     1260
ctcttcgctg tgctctggtg ccgggacatc accagtgcgc tggacacggc gaaacgcctg     1320
caggctcggt cggtaatgat caacgactgg ttcggtgagc tgccgatgac tccgcacgga     1380
ggccacaagc aaagcggcac cggacgcgag gaaggcctcg aagcgggtaca cggctacaca     1440
caggctcaagc acatcggcat caacctcgag ccgtcgcgccg caaagtccgc cgattgggcc     1500
ggtgcacctc tgtga

```

<210> 2

25 <211> 504

<212> PRT

<213> Mycobacterium austroafricanum

ES 2 392 742 T3

<400> 2

Met Thr Arg Thr Leu Ser Ala Asp Ala Asp Thr Arg Thr Ala Thr Pro
 1 5 10 15

Pro Leu Met Tyr Val Asn Gly Glu Trp Leu Pro Ala Arg Ser Gly Ala
 20 25 30

Thr Phe Pro Thr Ile Glu Pro Ser Thr Gly Arg Pro Ile Thr Glu Ile
 35 40 45

Pro Arg Gly Asp Ser Ser Asp Val Asp Ala Ala Val Lys Ala Ala Ala
 50 55 60

Asp Val Ala Val Glu Trp Gln Phe Thr Asp Ala Ile Thr Arg Ala Ala
 65 70 75 80

Leu Leu Arg Arg Leu Ala Glu Leu Val Ala Glu Asn Ala Glu Glu Leu
 85 90 95

Ala Arg Ile Glu Ser Leu Asp Ser Gly His Tyr Leu Ala Lys Ala Arg
 100 105 110

Glu Leu Val Thr Ala Ile Pro Leu Trp Leu Glu Tyr Trp Ala Gly Ala
 115 120 125

Ala Asp Lys Val Gly Gly Arg Thr Ile Ala Val Pro Gly Asn Lys Leu
 130 135 140

Ser Phe Thr Leu Leu Glu Pro Leu Gly Val Thr Ala His Ile Ile Pro
 145 150 155 160

Trp Asn Tyr Pro Leu Leu Ile Leu Ala Arg Ser Ile Ala Pro Ala Leu
 165 170 175

Ala Leu Gly Asn Thr Cys Val Val Lys Pro Ala Glu Asp Thr Ser Leu
 180 185 190

Ser Ala Leu Lys Phe Ala Glu Leu Val His Ala Ala Gly Phe Pro Ala
 195 200 205

Gly Val Phe Asn Val Val Thr Gly Tyr Gly Ser Glu Ala Gly Ala Ala
 210 215 220

ES 2 392 742 T3

Leu Ala Ala His Pro Glu Val Arg Gly Ile Thr Phe Thr Gly Ser Thr
225 230 235 240

Glu Thr Gly Arg Glu Ile Ala Arg Leu Gly Gly Gln His Ile Ala Gln
245 250 255

Val Asn Leu Glu Leu Gly Gly Lys Ser Pro Leu Val Val Phe Pro Asp
260 265 270

Ala Pro Leu Glu Asp Ala Val Glu Val Ala Val Gln Gly Phe Cys Ser
275 280 285

Arg Ala Gly Gln Val Cys Val Ala Gly Ser Arg Leu Phe Leu His Glu
290 295 300

Asp Ile Ala Asp Arg Phe Leu Glu Met Leu Val Ser Arg Leu Glu Thr
305 310 315 320

Val Thr Val Gly Asp Pro Phe Asp Gly Ala Thr Gln Met Gly Pro Leu
325 330 335

Ala Ser Lys Lys His Tyr Asp Arg Val Arg Glu Tyr Ile Glu Val Gly
340 345 350

Lys Gln Glu Ala Thr Leu Leu Tyr Gly Gly Gly Arg Pro Thr Asp Thr
355 360 365

Pro Asp Asp Gly Phe Phe Val Glu Pro Thr Val Phe Val Asp Val Ala
370 375 380

Thr Asp Ala Arg Ile Ala Arg Glu Glu Ile Phe Gly Pro Val Thr Ala
385 390 395 400

Val Met Arg Trp Ser Ser Val Asp Asp Leu Ile Ala Thr Ile Asn Asp
405 410 415

Ser Glu Phe Gly Leu Phe Ala Val Leu Trp Cys Arg Asp Ile Thr Ser
420 425 430

Ala Leu Asp Thr Ala Lys Arg Leu Gln Val Gly Ser Val Met Ile Asn
435 440 445

Asp Trp Phe Gly Glu Leu Pro Met Thr Pro His Gly Gly His Lys Gln
450 455 460

Ser Gly Thr Gly Arg Glu Glu Gly Leu Glu Ala Val His Gly Tyr Thr
465 470 475 480

Gln Val Lys His Ile Gly Ile Asn Leu Glu Pro Ser Pro Ala Lys Ser
485 490 495

Ala Asp Trp Ala Gly Ala Pro Leu
500

<210> 3
<211> 648

ES 2 392 742 T3

<212> DNA
 <213> Mycobacterium austroafricanum

<400> 3
 atgcctcaag aagacgcctt ggtgagcgcc ccattcgggg cgccgctcctt gggggaagat 60
 cgccccgatc aacccattt cgagaccggc cgaccgagac tcggcggact ggtcgcccc 120
 gtcgctttg atgagctcga ggcagcggc atcgacgccc tggccgacac gatgatcccc 180
 gccgaaggcg gctttccggc cgcaagcgac gtgggaatcg tcgatttctt cggccgctac 240
 acgactccca ccgattccg cgcgaagcac ttcccctacc tcgaagagga caagttgaag 300
 agcgcactcg cggggctcgg cgaagaattc gtcaacgccc acaccgatac ggcaccccag 360
 gcggtcctcc gattggagaa ggacgatgag gaggctctcg cgcagggtgag aagcctcgtg 420
 tattacggct actactccgc gaacgcagtg accgctgcca ttcaccagca gattccggcc 480
 gggcgcgact accacggacc cccactcccc tacggctatc tgcattgcat cgaggactgg 540
 gatgaagcgg cgctctccac atcggggcag ggctcaggct atgtcgccac cgacgatgtg 600
 5 gtccgagtgg atctcagcaa actcacctgg ctgaacaaca agacttga 648

<210> 4
 <211> 215
 <212> PRT
 10 <213> Mycobacterium austroafricanum

<400> 4
 Met Pro Gln Glu Asp Ala Leu Val Ser Ala Pro Phe Gly Ala Pro Ser
 1 5 10 15
 Leu Gly Glu Asp Arg Pro Asp Gln Pro His Phe Glu Thr Gly Arg Pro
 20 25 30
 Arg Leu Gly Gly Leu Val Ala Pro Val Ala Phe Asp Glu Leu Glu Ala
 35 40 45
 Ala Val Ile Asp Ala Leu Ala Asp Thr Met Ile Pro Ala Glu Gly Gly
 50 55 60
 Phe Pro Ala Ala Ser Asp Val Gly Ile Val Asp Phe Phe Gly Arg Tyr
 65 70 75 80
 Thr Thr Pro Thr Gly Phe Arg Ala Lys His Phe Pro Tyr Leu Glu Glu
 85 90 95
 Asp Lys Leu Lys Ser Ala Leu Ala Gly Leu Gly Glu Glu Phe Val Asn
 100 105 110

Ala Asp Thr Asp Thr Arg Thr Gln Ala Val Leu Arg Leu Glu Lys Asp
 115 120 125

Asp Glu Glu Phe Phe Ala Gln Val Arg Ser Leu Val Tyr Tyr Gly Tyr
 130 135 140

Tyr Ser Ala Asn Ala Val Thr Val Ala Ile His Gln Gln Ile Pro Ala
 145 150 155 160

Gly Arg Asp Tyr His Gly Pro Pro Leu Pro Tyr Gly Tyr Leu His Cys
 165 170 175

Ile Glu Asp Trp Asp Glu Ala Ala Leu Ser Thr Ser Gly Gln Gly Ser
 180 185 190

Gly Tyr Val Ala Thr Asp Asp Val Val Arg Val Asp Leu Ser Lys Leu
 195 200 205

Thr Trp Leu Asn Asn Lys Thr
 210 215

<210> 5

<211> 1659

5 <212> DNA

<213> Mycobacterium austroafricanum

<400> 5

gtgacaacat cgcgatca gaccgacgtt ctggatcatcg gttcgggacc gggcggcgcg 60
 ggcgtcacgc tcaagctggt gcaggccgga tacaaggtga cctgcctgga gcagggcct 120
 tgggtgacac cgcccagca cccgcattac caccgggaat gggagatcga aaagcaacgc 180
 ggatgggcct acgaccgaa cgtccgtgga ctcccggaag actaccgggt gaccggcttc 240
 accacgcctt atctcatgaa caacgtgggc ggtagcacga tgcactacgc cggccactgg 300
 ccgcgctaca agccggtcga cttccgcaag ggcaccgagc acggggtgga aggcacgatc 360
 gactggccga tcagctacga agagctcgcg ccgtactacg acgagaacga cgcgatctac 420
 ggcattctccg gcatgggtgg cgatccgtcg tatccggatc gaaccggagt cgaccgcgat 480
 ccaccgggtca aaccgggcaa gctggggcgc aacttcgctc aggcgctggg cgacctgggt 540
 tggcactggt ggccatcga caacgcgatc atcactcggc cacgcgaagg ccgcgaggct 600
 gacatcgccg caggcaacga gctctcgggt agcccagcgg gatcgtcag cacgccgacg 660
 cacaccact ggccgaccgc catcgcgctc ggagcggact tgcgtacca cgcccagatc 720
 gaacagatcc acacgaagaa cggcaaggcc accggtgcga cctacatcga caccgtacc 780
 ggcgcacggc acgagatcaa cgcgaagatc gtggtggtct cggccagcgg gatcggaaacc 840
 ccgagactgc tcctcatgag cgcgcagaag gggcatcccg acggtctggc caacagcaac 900
 ggtctcgtcg gcaaatacct gatgcaccac attcttcgcg ttctggcgag cgtggttcgc 960
 acaagccgga tggaaaggcta caagggagcc ttcggcgctc cactgtattc gcacgagttc 1020
 taccacaccg acaccaatcg cggcttcgct aacggtttcg gcatgcaggt ggcacgcagc 1080
 ttcggcgctg catacacagc aatgggcagc cacaccggtt acgtggcccc ctggggcaaa 1140
 tcgcatcga agttcttcaa cgaacacttc ggcaatcact tgatggtttt catgttcggc 1200
 gaagacctcc ccgtcgagac gaactgcgtg aactcgtatc ctgacgcaa agactcgagt 1260
 ggctccctg cggcgcgtgt caactgggaa ccacacgaga acgacatcgc gctggccaat 1320
 tacggcatcg accggatctt cgaggccgcg cgagcgttgg gcgctgtcga gaccaacgac 1380
 accggcgtgc tcaatcctcc gcccggttgg cacttgatgg gcacctgtcg gatgggtaac 1440
 aaccagaag attcggtcac caacaagtgg catcagacct gggatgtgcc gaatctcttc 1500
 gttgtcgacg ggagctcgt caccaccggc ggagctgtca acccgacatc gacgatcggc 1560
 gcgctcgccg tgcgggcagg agattacatc tcccgcgat tctccgacat cgtcgatcag 1620
 cgcaccacgc cgagcaacga agacgcacct gccatctaa 1659

5 <210> 6
 <211> 552
 <212> PRT
 <213> Mycobacterium austroafricanum

10 <400> 6

Met Thr Thr Ser Ala Asp Gln Thr Asp Val Leu Val Ile Gly Ser Gly
 1 5 10 15

Pro Gly Gly Ala Gly Val Thr Leu Lys Leu Val Gln Ala Gly Tyr Lys
 20 25 30

Val Thr Cys Leu Glu Gln Gly Pro Trp Val Thr Pro Pro Glu His Pro
 35 40 45

His Tyr His Arg Glu Trp Glu Ile Glu Lys Gln Arg Gly Trp Ala Tyr
 50 55 60

Asp Pro Asn Val Arg Gly Leu Pro Glu Asp Tyr Pro Val Thr Gly Phe
 65 70 75 80

Thr Thr Pro Tyr Leu Met Asn Asn Val Gly Gly Ser Thr Met His Tyr
 85 90 95

Ala Gly His Trp Pro Arg Tyr Lys Pro Val Asp Phe Arg Lys Gly Thr
 100 105 110

Glu His Gly Leu Glu Gly Thr Ile Asp Trp Pro Ile Ser Tyr Glu Glu
 115 120 125

Leu Ala Pro Tyr Tyr Asp Glu Asn Asp Ala Ile Tyr Gly Ile Ser Gly
 130 135 140

Met Val Gly Asp Pro Ser Tyr Pro Asp Arg Thr Gly Val Asp Arg Asp
 145 150 155 160

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

Glu Asn Asp Ile Ala Leu Ala Asn Tyr Gly Ile Asp Arg Ile Phe Glu
 435 440 445

Ala Ala Arg Ala Leu Gly Ala Val Glu Thr Asn Asp Thr Gly Val Leu
 450 455 460

Asn Pro Pro Pro Gly Trp His Leu Met Gly Thr Cys Arg Met Gly Asn
 465 470 475 480

Asn Pro Glu Asp Ser Val Thr Asn Lys Trp His Gln Thr Trp Asp Val
 485 490 495

Pro Asn Leu Phe Val Val Asp Gly Ser Ser Leu Thr Thr Gly Gly Ala
 500 505 510

Val Asn Pro Thr Ser Thr Ile Gly Ala Leu Ala Val Arg Ala Gly Asp
 515 520 525

Tyr Ile Ser Arg Arg Phe Ser Asp Ile Val Asp Gln Arg Thr Thr Pro
 530 535 540

Ser Asn Glu Asp Ala Pro Ala Ile
 545 550

<210> 7
 <211> 672
 <212> DNA
 <213> Mycobacterium austroafricanum

5

<400> 7
 atgtcggcaa ccgacgtccc gccgcttccc caatccatcg gccactgca cccggacggt 60
 gaagccggca gcttgtccga accaggtctc gagaccgcgg cgggccgctg ggtagctttc 120
 ggctggcca atctggcggg ggtgctggca gtgtccttgg cgggggtgga cctgctcgcc 180
 gacccccgac tcageccgtg gtccttctat ccaactcccgt tcaacgcggc tctgttctgg 240
 gccatcctct tcgttgtgtt catcggttc aatgccggat tcgcagggtt catccgactt 300
 tcacagccat ggcggggact cgccatcacc gtcgccacgg gcaccttcgc ggtcgccgtg 360
 acgtgggtgc tggcagccgg actcggcagt gtgaacgctg atttcgctgc cggccgagac 420
 ggcggcctcg gctacttcac cggcgcgctg ttcgtgctct tcgggttcgg taccttcgtg 480
 atagtcgtcc tcaattggca gactggcca tggccccaac tcggtctctc gcagcccggc 540
 gtgggattgg ccgagatcgc cgcggttgcc ggcccgacca tgctgctgta cttcgccttc 600
 ggcttgcccg cagtcagcgc aggcgggtgcc gaaccctgtg tggaattgga caccctcatg 660
 ggctggttct ag 672

10 <210> 8
 <211> 223
 <212> PRT
 <213> Mycobacterium austroafricanum

<400> 8

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

<210> 9

<211> 1233

<212> DNA

<213> Mycobacterium austroafricanum

<400> 9

tcagtcttgt tgacctatag cgctcagtgt tgtgcgcagt gccgcactg cgggcgctac 60

cgatggcgcc aacaaacctg ccatgttcgc tggatagcgg acggtcacga cggcgggttc 120
 ggcgaaccga tagggctcgac ccaccacgat gcccgcgaat tgctttctga gtcgggacat 180
 ctccactcga acggctccaa cgcgtgtggc gtctccgtac aggtctgccg caagcgcagg 240
 cgcagagcgg ccgtcagcgc tgtgggccag tatgagcagt atctcggcgt gccgcagtga 300
 aatgtcgtgc cgccagttac cgaatcggcc ctccatctcg agcgcaggag aggcactgtc 360
 gtgcagattc aacgtgacc gtgtgactgc tgactccgtg tcgcttccca ccagtcgcac 420
 caaccacccg ccgggtagca tctcgacctc gcaatcgccg agtgacggca ccataagtg 480
 gccgggcgcg atatcggcgg gcagcatgat ccgcctctgc acggtcaacg agtccactgc 540
 ggccacccaa ccgtcccat ccacggccac ggcaggcgac ccgactcgcg ccagtatcgg 600
 agcagcgacc gagcgcagcc ggttcagtgc gagttcgtgc tgcccacgca actccgattc 660
 ggccaaccgt gccacggcgt cgatgagggc gaccgtcacc ggatgcacgg tcgccgccgg 720
 tccggtcacg tcgaccacac cgatcacgcg gccggtgccg ggactacgga tcggtgcgcc 780
 cgcgcacgtc cacgggtggt ggctgcgggt gtaatgtcc gactgaagg tctggacggc 840
 gcgctgtgag acgagagcgg taccgagggc gttcgtaccg accgtcgttt cggccactg 900
 ggctccttcg acgaacccaa gcctgtccgc gttcgcacgc acacgagggc agccggatcg 960
 ccacagcaca gcgcactcgg catcggcgac ggcaacgata ctgtcgcccg tggacgtcaa 1020
 cggggcgagg ccgcgcgtca actccccgac caccgaagcc aagccggatt cggacctcaa 1080
 cacctcgacc gtgtccgact ccatcacggg cggcggtcga cggtcgggtc gcaaaccctt 1140
 ggccatgagt cgctgccagg aatcccagat gacacgacgt ggacgggccg gagcttgccc 1200
 gcccgccatc gtggcctcgt agacagccgc cat 1233

<210> 10
 <211> 410
 <212> PRT
 <213> Mycobacterium austroafricanum

5

<400> 10
 Met Ala Ala Val Tyr Glu Ala Thr Met Ala Gly Gly Gln Ala Pro Ala
 1 5 10 15
 Arg Pro Arg Arg Val Ile Trp Asp Ser Trp Gln Arg Leu Met Ala Arg
 20 25 30
 Gly Leu Arg Pro Asp Arg Arg Pro Pro Pro Val Met Glu Ser Asp Thr
 35 40 45
 Val Glu Val Leu Arg Ser Glu Ser Gly Leu Ala Ser Val Val Gly Glu
 50 55 60
 Leu Thr Arg Gly Leu Ala Pro Leu Thr Ser Thr Gly Asp Ser Ile Val
 65 70 75 80

10

Ala Val Ala Asp Ala₈₅ Glu Cys Ala Val Leu₉₀ Trp Arg Ser Gly Ser₉₅ Pro
 Arg Val Leu Ala₁₀₀ Asn Ala Asp Arg Leu₁₀₅ Gly Phe Val Glu Gly₁₁₀ Ala Gln
 Trp Ala Glu₁₁₅ Thr Thr Val Gly Thr₁₂₀ Asn Ala Leu Gly Thr₁₂₅ Ala Leu Val
 Ser Gln₁₃₀ Arg Ala Val Gln Thr₁₃₅ Phe Ser Ala Glu His₁₄₀ Tyr Asn Arg Ser
 Gln His₁₄₅ Pro Trp Thr Cys₁₅₀ Ala Gly Ala Pro Ile₁₅₅ Arg Ser Pro Arg Thr₁₆₀
 Gly Arg Val Ile Gly₁₆₅ Val Val Asp Val Thr₁₇₀ Gly Pro Ala Ala Thr₁₇₅ Val
 His Pro Val Thr₁₈₀ Val Ala Leu Ile Asp₁₈₅ Ala Val Ala Arg Leu₁₉₀ Ala Glu
 Ser Glu Leu₁₉₅ Arg Gly Gln His Glu₂₀₀ Leu Ala Leu Asn Arg₂₀₅ Leu Arg Ser
 Val Ala Ala Pro Ile Leu Ala₂₁₅ Arg Val Gly Ser Pro₂₂₀ Ala Val Ala Val
 Asp₂₂₅ Gly Asp Gly Trp Val₂₃₀ Ala Ala Val Asp Ser₂₃₅ Leu Thr Val Gln Arg₂₄₀
 Arg Ile Met Leu₂₄₅ Pro Ala Asp Ile Ala Pro₂₅₀ Gly His Leu Trp Val₂₅₅ Pro
 Ser Leu Gly Asp₂₆₀ Cys Glu Val Glu Met₂₆₅ Leu Pro Gly Gly Trp₂₇₀ Leu Val
 Arg Leu Val₂₇₅ Gly Ser Asp Thr Glu₂₈₀ Ser Ala Val Thr Arg₂₈₅ Val Thr Leu
 Asn Leu₂₉₀ His Asp Ser Ala Ser₂₉₅ Pro Ala Leu Glu Met₃₀₀ Glu Gly Arg Phe
 Gly Asn Trp Arg His Asp₃₁₀ Ile Ser Leu Arg His₃₁₅ Ala Glu Ile Leu Leu₃₂₀
 Ile Leu Ala His Ser₃₂₅ Ala Asp Gly Arg Ser Ala Pro Ala Leu Ala Ala
 Asp Leu Tyr Gly₃₄₀ Asp Ala Thr Arg Val₃₄₅ Gly Ala Val Arg Val₃₅₀ Glu Met

Ser Arg Leu Arg Lys Gln Phe Ala Gly Ile Val Val Gly Arg Pro Tyr
 355 360 365

Arg Phe Ala Glu Pro Ala Val Val Thr Val Arg Tyr Pro Ala Asn Met
 370 375 380

Ala Gly Leu Leu Ala Pro Ser Val Ala Pro Ala Val Arg Ala Leu Arg
 385 390 395 400

Thr Thr Leu Ser Ala Ile Gly Gln Gln Asp
 405 410

<210> 11

<211> 10327

<212> DNA

<213> Mycobacterium austroafricanum

5

<400> 11

tgcagcacag caccgcacc ccaccggcgg cgatacctcg ccaccggct caccgcgcg 60
 gtaggcacc accgcaaca cacgggcaga gcggtccagc acctgtggag aatcatcgac 120
 accctcaccg actaccttcg cgccgtcgac agcggtcacc gcacctacac cgaagcgttc 180
 cacgaactgc tgccctacta cgccgacgca tcacaagccg ttgggactgc gaccaatccg 240
 gcctgacgac atgaccaacc cgggacatat gggcgccgaa tcgggcgttt cgccgacagc 300
 ccacggcacc ctcaaggcgt ccttctgacc caaccgaagt gcagcctatt cacctgaatg 360
 acagcggcgc ctattgttg acacatggtg tcggctacac ctcacgagtc ccgctcggtc 420
 actgcgcaac gctcgtcggg tgacgaacag cgcctcgacg ccgctatgc ggaactggcc 480
 acgtcgttca acgagaagt caccgacgcc gagagacgcg cggcgcgcac caagtacatc 540
 gaccgccgcg aaccggccg tcaatgaatg cggcgcgcac cttgtcgtcg gccttctcat 600
 atcaagtatc gaatggtcac acagggcgcg gcgaccggtg gggccgtccc atgggcactt 660
 cccggcttcc ttcgaccact cgccacgggg ccggcaggac gtcgcctcga cgcgccaact 720
 caaatgacgc tgccaatcac aagccatct caccgacgct ggcggtccta tgcacgggt 780
 caccaggtcg actgagaacg gacaaagtgg gtccggattg gtcgacgaac tgcgcgaggc 840
 cggctgcgcc gtgggtgtag tgaagggccg caccacggc gctgatcaag acagaccgta 900
 gcggcgtcgc agctcctcgc cgctgatcgt gttccccgcc gcataatccg catccgcggc 960
 ggccagttca tcgactaacc gctcattggc agcatggatt gcatcagcca atgctgcgaa 1020
 gtcggccggc gtgctgccag caggatcgtc actcatggct ccattctcgt cacaaccgac 1080
 tcgcccgtgg cgcgacggat ggacccgagc cgcggcgatc cgctccaact cttttagcga 1140
 cagcacgtag tccgctagtt agtattggcg gtcgcgctgg cctcgatgcc gcggtgttcg 1200
 ccgagcatcc gctttaggt gtgttacggt ttggcgattt gtctactgt tttttagatc 1260
 tctgtttcga cctgacccg ctcttgaatt tgttcggtaa gcgatcggtc ggccaactgg 1320
 gtgtttctgt tcgcatgtct gcatattggt cctattcatg caccactgat tactcttgtg 1380

10

gagaataagc gtgtcgggcc caacatgtat tatctgaacg gtgagtagag aaagagcggc 1440
gcttaagcct tgctgactac agcgtgcaa agacgcctct agtacgcgcg attgcttaac 1500
cggaaaccgta cccagccgcg caccgcacg gcagtcgcc acggcggctc cctgaacaat 1560
caccaccg aggctgaaaa gctccgggag tcgataaaca gatttgtcag tcagtccaag 1620
ctattcgaag gcaacgctgt gcctagtctc gctgatactt tggcgaaagat ccaacagaca 1680
acgatgcaaa gcttgcccc gatctacgtc ccgcagctgc ccgaatcgat gttcgacagc 1740
cttcgcaaga ccttaaaggg catctttccc gctaattggc cgagcgaatt gcctgacttc 1800
gaccgcatcg aggaggctcat agagaaggac ggcattccca tcgtccacat cccgcgagct 1860
gagatcgtcc aagagatctg cgacgctcac gactacgacg cacgcgtgca aatcttggctc 1920
gaccgtgcaa gcgacatcgc tgacgattgc cgagatgacg ttcagcgtgt ctacgacgag 1980
atagtggata agcagctccc cctagcccag cgagcagctg aggcatacca ggctggatac 2040
ttcaggtctg cccaggcatt ggcgggtgagt gtgtgagata cctatctgaa gaagatgtat 2100
caaggcacgc gctatgcgga catacaaagt gatctcacgc tcgataactc gaacgaggtg 2160
gccgtcggc tgccctcaa ctttcacttc gcattggcat ccgcagttcc gttcctcacc 2220
ccctggaacc ctgcgagggg tgaaaagccg cccacgaggc tgtcacgac gcgagtatcc 2280
atgacgcaag caccgacct atgacgctgc tgaatgcgac gattgcgata atgctcctga 2340
ccagcatgac tgttgctatc gactacatgg ctcggcgcgt gcggaagaac aaggagatt 2400
gcgacacctg aagtcctcgc acgccgcgtg acaggtacgc ctcctctatc cacgtcggcg 2460
aaggcgatcg ctaaccgcg attgcggact tggcaacatc atcaagacgc cctccttcac 2520
ctccgtctca atgcagttg catcctcaca ggtcgcgagt gacacctta ctgagaccgg 2580
acaactccgc gcgccgcggc caaccacacc accatcaccg acatcgaagc cgccctgcgc 2640
gaccaacggc tcctgaacc cgcgccgca accatcggc acgccgacc ctatggcgcc 2700
agctcgggtg ctggccacct cacgcgtaca tcacgccggc gactcccgca tcggacgctt 2760
aagtcagtct tgttgacct tagcgtcag tgttgtgcgc agtgcccgca ctgcgggccc 2820
taccgatggc gccaaacaac ctgccatggt cgctggatag cggacggtca cgacggcggg 2880
ttcggcgaac cgatagggtc gaccaccac gatgcccgcg aattgctttc tgagtcggga 2940
catctccact cgaacggctc caacgcgtgt ggcgtctccg tacaggtctg ccgcaagcgc 3000
aggcgcagag cggccgtcag cgctgtgggc cagtatgagc agtatctcgg cgtgccgacg 3060
tgaaatgtcg tgccgccagt taccgaatcg gccctccatc tcgagcgcag gagaggcact 3120
gtcgtgcaga ttcaacgtga cccgtgtgac tgctgactcc gtgtcgcttc ccaccagtcg 3180
caccaaccac ccgccgggta gcatctcgac ctcgcaatcg ccgagtgcg gcaccataa 3240
gtggccgggc gcgatatcgg cgggcagcat gatccgcctc tgcacggtca acgagtccac 3300
tgccgccacc caaccgtccc catccacggc cacggcaggc gaccgcactc gcgccagtat 3360
cggagcagcg accgagcgcg gccggttcag tgcgagttcg tgctgccac gcaactccga 3420

ttcggccaac cgtgccacgg cgtc gatgag ggcgaccgtc accggatgca cggtcgccgc 3480
 cgggtccggtc acgtcgacca caccgatcac gcgggccggtg cgcggactac ggatcgggtg 3540
 gcccgcgcac gtccacgggt gttggctgcg gttgtaatgc tccgcactga aggtctggac 3600
 ggcgcgctgt gagacgagag cgttaccgag ggcgttcgta ccgaccgtcg tttcggccca 3660
 ctgggctcct tcgacgaacc caagcctgtc cgcgttcgcc agcacacgag gcgagccgga 3720
 tcgccacagc acagcgcact cggcatcggc gacggcaacg atactgtcgc ccgtggacgt 3780
 caacggggcg aggccgcgcg tcaactcccc gaccaccgaa gccaagccgg attcggacct 3840
 caacacctcg accgtgtccg actccatcac gggcggcggt cgacggtcgg gtcgcaaacc 3900
 cctggccatg agtcgctgcc aggaatccca gatgacacga cgtggacggg ccggagcttg 3960
 cccgcccgc atcgtggcct cgtagacagc cgccatcagc cgcgagtagt cgcgcggatg 4020
 cctgtccggg cgcgatcgcg ggctccagggt cgggcagtgc tgtgcatcac ctccatcatg 4080
 ccctaccgtg ctgcggctct accacaggaa ccgccagggt acgcctcgaa ttcaggagta 4140
 cgcccggcgc tgtaacgccc gtgtaactct tgcgggcgca cccctcgtga atgtgtgctg 4200
 cttcacacga acggccgtac atgccgctcg tcacagcggg cggaacagac agtgacctgc 4260
 ctgtccggcg gccgggccc cctcggtgaa tcaccgacct ggcacgcgag tcattcatag 4320
 gagcgagaac aacatgacct gaactctctc ggctgatgcc gacacccgca cggcgacacc 4380
 tccgctgatg tacgtcaacg gcgagtggct gcccgccgc agtggggcca ctttcccac 4440
 catcgaacct agcacgggtc gaccgatcac tgagattccc cgcggggact cgagcgatgt 4500
 ggacgcggcg gtgaaagccg ctgccgacgt ggccggtgag tggcagttca ccgatgccat 4560
 caccgcgcc gccctgctca ggcgattggc ggagctgggt gcagagaacg ccgaggagct 4620
 ggcgcggatc gagtcgctgg actcgggtca ctatctggcg aaggcgcgtg aactggtgac 4680
 cgcgataccc ctgtggctcg agtactgggc cggcgcagct gacaaagtgg gcggccgcac 4740
 catcgtgta ccgggtaaca aactcagctt caccttggtg gaaccgctgg gcgtcaccgc 4800
 gcacatcatc ccgtggaact acccgctggt gatccttgct cgggtccatc ccccggcact 4860
 cgcattgggc aacacctgtg tcgtcaagcc cgctgaggac acgtccctgt cagcgtgaa 4920
 gttcggcgag ctggtacacg ccgccggtt tcccgcggga gtgttcaatg tggtgaccgg 4980
 ttacggttcc gaagccggcg cggctcttg cgtcacccc gaggtgcgcg gaatcacctt 5040
 caccggttcg accgagaccg ggcgggagat cggccgactg ggcggccaac acatcggcca 5100
 ggtcaacctg gaactcggcg ggaagagccc gttggtcgtc tttcccgcg cgcgctcga 5160
 agacgccgta gaggtggccg tacagggctt ctgctcacgg gcagggaag tgtgtgtcgc 5220
 cgggagccgg ctcttctcc atgaggacat cggcagccgg ttctcgcgaga tgctcgtttc 5280
 gcgactcgag actgtcaccg tcggcgacct gttcgcgggt gcgaccaga tgggtccgct 5340
 cgctcgaag aagcactacg accgtgtgcg tgagtacatc gaggtcggga agcaggaggc 5400
 gaccctgctc tacggcggcg gtcggccgac ggacacgccc gatgacgggt tcttcgtcga 5460

gccaacggtt ttcgtcgacg tcgcaacgga tgcgcggatc gcacgcgagg agatcttcgg 5520
 gcccgtcaca gcggtgatgc ggtggtcatc ggtcgacgat ctgatcgcca ccatcaatga 5580
 ttcggaattc ggtctcttcg ctgtgctctg gtgccgggac atcaccagtg cgctggacac 5640
 ggcgaaacgc ctgcaggtcg gctcggtaat gatcaacgac tggttcgggtg agctgccgat 5700
 gactccgcac ggaggccaca agcaaagcgg caccggacgc gaggaaggcc tcgaagcgg 5760
 acacggctac acacaggtca agcacatcgg catcaacctc gagccgtcgc ccgcaaagtc 5820
 cgccgattgg gccggtgcac ctctgtgatc gatcggccgc ttgccgtgcg gtgctgaaca 5880
 acgcattcca agaacaacac acatgaagga gagtcaatgc ctcaagaaga cgccttggtg 5940
 agcgcctcat tcggggcgcc gtccttgggg gaagatcgtc ccgatcaacc ccatttcgag 6000
 accggccgac cgagactcgg cggactggtc gccccgctc cgtttgatga gctcgaggca 6060
 gcggtcatcg acgccctggc cgacacgatg atccccgccg aaggcggctt tccggccgca 6120
 agcgacgtgg gaatcgtcga tttcttcggc cgctacacga ctcccaccgg attccgcgcg 6180
 aagcacttcc cctacctcga agaggacaag ttgaagagcg cactcgcggg gctcggcgaa 6240
 gaattcgtca acgccgacac cgatacgcgc acccaggcgg tcctccgatt ggagaaggac 6300
 gatgaggagt tcttcgcgca ggtgagaagc ctcgtgtatt acggctacta ctccgcgaac 6360
 gcagtgaccg tcgccattca ccagcagatt ccggccgggc gcgactacca cggacccccca 6420
 ctcccctacg gctatctgca ttgcatcgag gactgggatg aagcggcgct ctccacatcg 6480
 gggcagggct caggctatgt cgccaccgac gatgtggtcc gagtggatct cagcaaactc 6540
 acctggtga acaacaagac ttgaacacaa gacttgacga ggagcacaag tgacaacatc 6600
 cgccgatcag accgacgttc tggatcatcg ttcgggaccg ggcggcgcgg gcgtcacgct 6660
 caagctggtg caggccggat acaaggtgac ctgcctggag caggggcctt ggggtgacacc 6720
 gcccgagcac ccgattacc accgggaatg ggagatcgaa aagcaacgcg gatgggccta 6780
 cgacccgaac gtccgtggac tcccggaaga ctaccgggtg accggcttca ccacgcctta 6840
 tctcatgaac aacgtgggcg gtagcacgat gcactacgcc ggccactggc cgcgctacaa 6900
 gccggtcgac ttccgcaagg gcaccgagca cgggttgga ggcacgatcg actggccgat 6960
 cagctacgaa gagctcgcgc cgtactacga cgagaacgac gcgatctacg gcatctccgg 7020
 catggtgggc gatccgtcgt atccggatcg aaccggagtc gaccgcgatc caccggtcaa 7080
 accgggcaag ctggggcgca acttcgctca ggcgctgggc gacctgggtt ggcactggtg 7140
 gccatcggac aacgcgatca tcaactcggc acgcgaaggc cgcgaggctg acatcgccgc 7200
 aggcaacgag ctctcgggta gcccgacggg atcgctcagc acgccgacgc acaccactg 7260
 gccgaccgcc atcgcgctcg gagcggactt gcgtaccac gcccgagtcg aacagatcca 7320
 cacgaagaac ggcaaggcca ccggtgcgac ctacatcgac acccgtagcg gcgcacggca 7380
 cgagatcaac gcgaagatcg tgggtgctc ggccagcggg atcggaacct cgagactgct 7440
 cctcatgagc gcgcagaagg ggcattcccga cggctctggc aacagcaacg gtctcgtcgg 7500

ES 2 392 742 T3

caaatacctg atgcaccaca ttcttcgctg tctggcgagc gtggttcgca caagccggat 7560
 ggaaggctac aagggagcct tcggcgctcc actgtattcg cacgagttct accacaccga 7620
 caccaatcgc ggcttcgtca acggtttcgg catgcagggtg gcacgcagct tcggcgctgc 7680
 atacacagca atgggcagcc acaccggtta cgtggccccc tggggcaa at cgcatcgcaa 7740
 gttcttcaac gaacacttcg gcaatcactt gatggttttc atgttcggcg aagacctccc 7800
 cgtcgagacg aactgcgtga cactcgatcc tgacgccaaa gactcgagtg gcctccctgc 7860
 ggcgcgctgc aactgggaac cacacgagaa cgacatcgcg ctggccaatt acggcatcga 7920
 ccggatcttc gaggccgcgc gagcgttggg cgctgtcgag accaacgaca ccggcgctgc 7980
 caatcctccg cccggttggc acttgatggg cacctgtcgg atgggtaaca acccagaaga 8040
 ttcggtcacc aacaagtggc atcagacctg ggatgtgccc aatctcttcg ttgtcgacgg 8100
 gagctcgctc accaccggcg gagctgtcaa cccgacatcg acgatcggcg cgctcgccgt 8160
 gcgggcagga gattacatct cccgccgatt ctccgacatc gtcgatcagc gcaccacgcc 8220
 gagcaacgaa gacgcacctg ccatcta atc cattcgaact gccgatccac ctcccgcggg 8280
 cggcgggagg tggtatcggg tccatttgat ctcttcgagc ccacaggagc accccagatg 8340
 tcggcaaccg acgtcccgcc gcttcccaa tccatcggcc cactgcacc ggacgttgaa 8400
 gccggcagct tgtccgaacc aggtctcgag acccgggcg gccgctgggt agctttcggc 8460
 ctggccaatc tggcgggtgt gctggcagtg tccttggcgg ggtggtacct gctcgccgac 8520
 ccccgactca gcccggtgtc cttctatcca ctcccgttca acgcggtctt gttctgggcc 8580
 atcctcttcg ttgtgttcat cggcttcaat gccggattcg cagggttcat ccgactttca 8640
 cagccatggc ggggactcgc catcaccgtc gccacgggca tcttcgcggt cgccgtgacg 8700
 tgggtgctgg cagccggact cggcagtggt aacgctgatt tcgctgccgg ccgagacggc 8760
 ggcttcggct acttcaccgg cgcgctgttc gtgctcttcg ggttcggtac ctctgtgata 8820
 gtcgtcctca attggcagca ctggccatgg ccccaactcg gtctctcga gcccgcgctg 8880
 ggattggccg agatcgccgc ggttgccggc ccgaccatgc tgctgtactt cgcttcggc 8940
 ttgcccgcag tcagcgcagg cggtgccgaa cccgtgttg aattggacac cctcatgggc 9000
 tggttctagc ccgaatcttt cgcggaatgt tgctttggc ggtgtgtaga taagcgaag 9060
 tgctgtcct gcaagggata attggacttc tctacggtc aaagatcctg acgggaaggc 9120
 acttcgcagg tgaagaatat cgcggccgca tcgcggtgaa aagtgtcggc cgacggccat 9180
 ggctcgtgt cgcatgccgg gatgggcctg ctacgtgaac tggccgatcg gaccgggtta 9240
 tcggcgagg tcacggctgc tttggccgac acctaccgcg ggccgtgggt gtatgcgcc 9300
 ggagacgtgt tcgctgatct ggcggctgcg gttgctgacg gggcggactg catcgacggg 9360
 gtcggccagc tctcggcgca ccgtgagcac gccttcggtg cgaaagcctc gacgaccag 9420
 atgtggcggc tggtcgatga gcgcatcgac gccgcacacc taccgcgggt gcgggcggcc 9480
 cgagcgggcg cgcgggcggc ggcctgggcc gccggtgccc ccccgctcc gggtgactgg 9540

ES 2 392 742 T3

ctgcacatcg acatcgacgc caccctggtg atcgatcact cgcacaaca aggcggtgcg 9600
 acgccgacct ggaagaagtc gttcggggcac catccgctgc tggcatttct ggaccgcccc 9660
 gagatcgccg gcggggaagc cctggccggg ctgctacgca ccggcaacgc cggctccaac 9720
 accgccagcg accacgtcat cgtgctggcc caagcgtgag cagcgtgccc cgccgctggt 9780
 cgacccgacc ccagggtctg cgatcatccc gacaagccca aggtgctggt gcgctgagac 9840
 accgccggag ccacccacac cttcgccgac gcctgccggg ccgccggggt ggggttctcc 9900
 ttcggttacc ccgtcgactg gcgctgcca gacgcggtg acaccctcaa cctcgccgag 9960
 ggctggtacc cggcgatcgg caccgccggc ggtctccgag acggcgctg gatcgccgaa 10020
 gccaccaccc tggtaaacct gtcgctggtg ccaccgggca cgcggctgat cctgagcaaa 10080
 gaaaggccgc atccggggcg ccagctgca ttcaccgagc ccgacgggat gcggatcaca 10140
 gcgttcatca ccgacacacc acccggcgtc gtgtcaggac aggtcgccgg cctggaactg 10200
 cgccaccgtc agcacgcccg cgtcgaagac cgcattccgc aactcaaagc caccggcctg 10260
 cgcaacctgc cctgccactc attttgggccc aacgccgat ggctggaaat cgccctcacc 10320
 gccgccg 10327

<210> 12

<211> 1480

<212> DNA

<213> *Mycobacterium austroafricanum*

5

<400> 12

agagtttgat tctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgaac 60
 ggaaaggccc ttcgggttac tcgagtggcg aacgggtgag taacacgtgg gtgatctgcc 120
 ctgcactttg ggataagcct gggaaactgg gtctaatacc gaatacacc ttctggctgc 180
 atggtctggt tggggaaagc ttttgcggtg tgggatgggc ccgccccta tcagcttgtt 240
 ggtgaggtta cggctcacca aggcgacgac gggtagccgg cctgagaggg tgaccggcca 300
 cactgggact gagatacggc ccagactcct acgggaggca gcagtgggga atattgcaca 360
 atgggcgcaa gcctgatgca gcgacgccgc gtgagggatg acggccttcg ggttgtaaac 420
 ctctttcgcc agggacgaag cgcaagtgac ggtacctgga gaagaaggac cggccaacta 480
 cgtgccagca gccgcggtaa tacgtaggtt ccgagcgttg tccggaatta ctgggcgtaa 540
 agagctcgta ggtggtttgt cgcgttggtc gtgaaaactc acagctcaac tgtgggcgtg 600
 cgggcgatac gggcagactg gagtactgca ggggagactg gaattcctgg tgtagcggtg 660
 gaatgcgag atatcacgag gaacaccggt ggcgaaggcg ggtctctggg cagtaactga 720
 cgctgaggag cgaaagcgtg gggagcgaac aggattagat accctggtag tccacgccgt 780
 aaacggtggg tactaggtgt gggtttcctt ccttgggatc cgtgccgtag ctaacgcatt 840
 aagtacccc cctggggagt acggccgcaa ggctaaaact caaagaaatt gacgggggccc 900
 cgcaaacgcg gcggagcatg tggattaatt cgatgcaacg cgaagaacct tacctggggtt 960
 tgacatgcac aggacgtgcc tagagatagg tattcccttg tggcctgtgt gcagggtggtg 1020

10

catggctgtc gtcagctcgt gtcgtgagat gttggggttaa gtcccgaac gagcgcaacc 1080
 cttatcttat gttgccagcg cgtaatggcg gggactcgtg agagactgcc ggggtcaact 1140
 cggaggaagg tggggatgac gtcaagtcac catgcccctt atgtccaggg cttcacacacat 1200
 gctacaatgg ccggtacaaa gggctgcat gcccgtgaggt ggagcgaatc ctttcaaagc 1260
 cggctctcagt tccgatcggg gtctgcaact cgaccccgtg aagtcggagt cgctagtaat 1320
 cgcagatcag caacgctgcy gtgaatacgt tcccgggcct tgtacacacc gcccgctcacg 1380
 tcatgaaagt cggtaacacc cgaagccggt ggcctaacc cttgtgggag ggagccgctc 1440
 aagggtgggat cggcgattgg gacgaagtcg taacaaggta 1480

<210> 13
 <211> 1362
 <212> DNA
 <213> Mycobacterium austroafricanum

5

<400> 13

tccaacggaa aggcccttcg gggactcga gtggcgaac ggtgagtaac acgtgggtga 60
 tctgccctgc actttgggat aagcctggga aactgggtct aataccgaat acacccttct 120
 ggctgcatgg tctggttggg gaaagctttt gcggtgtggg atgggcccgc ggcctatcag 180
 cttgttggtg aggttacggc tcaccaaggc gacgacgggt agccggcctg agaggggtgac 240
 cggccacact gggactgaga tacggcccag actcctacgg gaggcagcag tggggaatat 300
 tgcacaatgg gcgcaagcct gatgcagcga cgccgctga gggatgacgg ccttcgggtt 360
 gtaaacctct ttcgccaggg acgaagcgca agtgacggta cctggagaag aaggaccggc 420
 caactacgtg ccagcagccg cggtaatacgt tagggtccga gcggtgtccg gaattactgg 480
 gcgtaaagag ctcgtaggtg gtttgtcgcg ttgttcgtga aaactcacag ctcaactgtg 540
 ggcgtgcggg cgatacgggc agactggagt actgcagggg agactggaat tcctggtgta 600
 gcggtggaat gcgcagatat caggaggaac accggtggcg aaggcgggtc tctgggcagt 660
 aactgacgct gaggagcga aagcgtgggga gcgaacagga ttagataccc tggtagtcca 720
 cgccgtaaac ggtgggtact aggtgtgggt ttccttcctt gggatccgtg ccgtagctaa 780
 cgattaagt accccgcctg ggggagtacg ccgccagggc taaaactcaa agaaattgac 840
 gggggcccgc acaagcggcg gagcatgtgg attaattcga tgcaacgcga agaaccttac 900
 ctgggtttga catgcacagg acgtgcctag agataggtat tcccttgtgg cctgtgtgca 960
 ggtggtgcat ggctgtcgtc agctcgtgtc gtgagatgtt gggttaagtc ccgcaacgag 1020
 cgcaaccctt atcttatgtt gccagcgcgt aatggcgggg actcgtgaga gactgccggg 1080
 gtcaactcgg aggaagggtg ggatgacgct aagtcacat gcccccttatg tccagggcct 1140
 cacacatgct acaatggccg gtacaaaggg ctgcatgccc gtgaggtgga gcgaatcctt 1200
 tcaaagccgg tctcagttcg gatcggggtc tgcaactcga ccccgtgaag tcggagtcgc 1260
 tagtaatcgc agatcagcaa cgctgcgggtg aatacgttcc cgggccttgt acacaccgcc 1320
 cgtcacgtca tgaaagtacg gtaacacccg aagccggtgg ta 1362

10

ES 2 392 742 T3

<210> 14
 <211> 591
 <212> DNA
 5 <213> Mycobacterium austroafricanum
 <400> 14

acgggtctcgt cggcaaatac ctgatgcacc acattcttcg cgttctggcg agcgtggttc 60
 gcacaagccg gatggaaggc tacaagggag ccttcggcgc tccactgtat tcgcacgagt 120
 tctaccacac cgacaccaat cgcggcttcg tcaacggttt cggcatgcag gtggcacgca 180
 gcttcggcgc tgcatacaca gcaatgggca gccacaccgg ttacgtggcc ccctggggca 240
 aatcgcacgc caagttcttc aacgaacact tcggcaatca cttgatggtt ttcattgttcg 300
 gcgaagacct ccccgtcgag acgaactgcg tgacactcga tcctgacgcc aaagactcga 360
 gtggcctccc tgcggcgcgt gtcaactggg aaccacacga gaacgacatc gcgctggcca 420
 attacggcat cgaccggatc ttcgaggccg cgcgagcgtt gggcgctgtc gagaccaacg 480
 acaccggcgt gctcaatcct ccgccgggtt ggcacttgat gggcacctgt cggatgggta 540
 acaaccaga agattcggtc accaacaagt ggcacagac ctgggatgtg c 591

10 <210> 15
 <211> 20
 <212> DNA
 <213> artificial

15 <220>
 <223> cebador Bott1

<400> 15
 tgcacacagg ccacaacca 20

20 <210> 16
 <211> 21
 <212> DNA
 <213> artificial

25 <220>
 <223> cebador Bott2

<400> 16
 30 gagagttga tcctggctca g 21

35 <210> 17
 <211> 20
 <212> DNA
 <213> artificial

<220>
 <223> cebador 244

<400> 17
 40 cccactgctg cctcccgtag 20

45 <210> 18
 <211> 20
 <212> DNA
 <213> artificial

ES 2 392 742 T3

<220>
 <223> cebador Tb11
 <400> 18
 5 accaacgatg ggtgtccat 20
 <210> 19
 <211> 20
 <212> DNA
 10 <213> artificial
 <220>
 <223> cebador Tb12
 15 <400> 19
 cttgtcgaac cgcatacct 20
 <210> 20
 <211> 24
 20 <212> DNA
 <213> artificial
 <220>
 <223> cebador MaFV2
 25 <400> 20
 gtctaatacc gaatacacc ttct 24
 <210> 21
 <211> 24
 30 <212> DNA
 <213> artificial
 <220>
 <223> cebador MaRV6
 35 <400> 21
 gtagtggcc ggtccttct ctcc 24
 <210> 22
 <211> 20
 <212> DNA
 <213> artificial
 45 <220>
 <223> cebador MF1
 <400> 22
 50 tgagaagcct cgtgtattac 20
 <210> 23
 <211> 18
 <212> DNA
 <213> artificial
 55 <220>
 <223> cebador MR1
 <400> 23
 60 gagataaggc gtggtgaa 18
 <210> 24
 <211> 20
 <212> DNA
 65 <213> artificial

<220>
 <223> cebador MF2
 5 <400> 24
 agtgacggca cccataagtg 20
 <210> 25
 <211> 21
 10 <212> DNA
 <213> artificial
 <220>
 <223> cebador MR2
 15 <400> 25
 tcgaggtgtt gaggtccgaa t 21
 <210> 26
 <211> 18
 20 <212> DNA
 <213> artificial
 <220>
 25 <223> cebador MF3
 <400> 26
 atcatcccgt ggaactac 18
 30 <210> 27
 <211> 18
 <212> DNA
 <213> artificial
 35 <220>
 <223> cebador MR3
 <400> 27
 40 tgacctgggc gatgtgtt 18
 <210> 28
 <211> 18
 <212> DNA
 <213> artificial
 45 <220>
 <223> cebador MF4
 <400> 28
 50 atcagacctg ggatgtgc 18
 <210> 29
 <211> 19
 <212> DNA
 55 <213> artificial
 <220>
 <223> cebador MR4
 60 <400> 29
 ggctgtgaaa gtcggatga 19
 <210> 30
 <211> 20
 <212> DNA
 65 <213> artificial

ES 2 392 742 T3

<220>
<223> cebador RT-PCR-F1

5 <400> 30
agtgacggca cccataagtg 20

<210> 31
<211> 21
10 <212> DNA
<213> artificial

<220>
<223> cebador RT-PCR-R1

15 <400> 31
tcgaggtgtt gaggtccgaa t 21

<210> 32
<211> 20
20 <212> DNA
<213> artificial

<220>
<223> cebador RT-PCR-F2

25 <400> 32
gcaggtcggc tcggtaatga 20

<210> 33
<211> 20
30 <212> DNA
<213> artificial

<220>
<223> cebador RT-PCR-R2

35 <400> 33
gtaatacacg aggcttctca 20

40 <210> 34
<211> 20
<212> DNA
<213> artificial

45 <220>
<223> cebador RT-PCR-F3

50 <400> 34
acggtctcgt cgcaaatatc 20

<210> 35
<211> 18
55 <212> DNA
<213> artificial

<220>
<223> cebador RT-PCR-R3

60 <400> 35
gcacatccca ggtctgat 18

REIVINDICACIONES

- 5 1. Polipéptido aislado o purificado que tiene una actividad en la vía de degradación del MTBE, y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionado del grupo que consiste en el alcohol *terc*-butílico (TBA), el 2-metil-1,2-propanodiol (2-M-1,2-PD), el hidroxisobutiraldehído y el ácido hidroxisobutírico (HIBA), seleccionándose dicho polipéptido del grupo que consiste en:
- a) un polipéptido que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en la SEQ ID NO:2 o la SEQ ID NO:10,
- b) un polipéptido que comprende una secuencia de aminoácidos que presenta al menos 70 % de identidad, preferentemente 75%, 80%, 90%, 95%, 98% o 99% de identidad con la secuencia de aminoácidos de un polipéptido que comprende la SEQ ID NO:2, el polipéptido tal como el definido en a) o b) de la secuencia SEQ ID NO:2 que tiene una actividad de deshidrogenación del hidroxisobutiraldehído a HIBA, y
- 10 el polipéptido tal como el definido en a) de la secuencia SEQ ID NO:10 que tiene una actividad de regulador transcripcional
- 15 2. Ácido nucleico purificado o aislado, caracterizado porque codifica al menos un polipéptido según la reivindicación 1.
3. Ácido nucleico aislado o purificado que codifica al menos un polipéptido que tiene una actividad en la vía de degradación del MTBE, o al menos uno de los intermedios catabólicos del MTBE, preferentemente el TBA, el 2-M-1,2-PD, el hidroxisobutiraldehído y el HIBA, seleccionado del grupo que consiste en :
- 20 e) un ácido nucleico que comprende al menos una cualquiera de las secuencias nucleotídicas seleccionada del grupo que consiste en la SEQ ID NO:1 o la SEQ ID NO:9, o un complemento de estas,
- f) un ácido nucleico que comprende al menos una secuencia nucleotídica que presenta al menos 70 % de identidad, preferentemente 75%, 80%, 90%, 95%, 98% o 99% de identidad con la secuencia nucleotídica de un ácido nucleico tal como el definido en e).
4. Ácido nucleico, caracterizado porque comprende la secuencia nucleotídica SEQ ID NO:11.
- 25 5. Ácido nucleico según una cualquiera de las reivindicaciones 2 a 4, caracterizado porque su transcripción está bajo el control de un promotor único.
6. Ácido nucleico según una cualquiera de las reivindicaciones 2 a 5, caracterizado porque su transcripción está inducida por la presencia de MTBE, y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionado del grupo que consiste en el TBA, el 2-M-1,2-PD, el hidroxisobutiraldehído y el HIBA.
- 30 7. Ácido nucleico según la reivindicación 6, caracterizado porque está organizado en operón.
8. Polipéptido según la reivindicación 1 o ácido nucleico según una cualquiera de las reivindicaciones 2 a 7, caracterizado porque es aislado y purificado de una cepa bacteriana elegida entre las cepas *Mycobacterium austroafricanum* registradas en la colección CNCM bajo los números I-3401 y I-2562.
9. Ácido nucleico recombinante, caracterizado porque comprende un ácido nucleico según una cualquiera de las reivindicaciones 2 a 8.
- 35 10. Vector caracterizado porque comprende un ácido nucleico según una cualquiera de las reivindicaciones 2 a 9.
11. Célula hospedante aislada que comprende bien al menos uno de los ácidos nucleicos según una cualquiera de las reivindicaciones 2 a 9, o bien al menos uno de los vectores según la reivindicación 10.
- 40 12. Procedimiento de identificación de una célula o de un ácido nucleico de una célula susceptible de degradar el MTBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionado del grupo que consiste en el TBA, el 2-M-1,2-PD, el hidroxisobutiraldehído y el HIBA, que comprende:
- opcionalmente una etapa de siembra de la célula sobre el medio suplementado en MTBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionado del grupo que consiste en el TBA, el 2-M-1,2-PD, el hidroxisobutiraldehído y el HIBA,
- 45 - una etapa de cribado de los ácidos nucleicos de dicha célula por hibridación con al menos una sonda, comprendiendo dicha sonda la secuencia nucleotídica complementaria de un fragmento de al menos 50 nucleótidos consecutivos de la secuencia nucleotídica seleccionada del grupo que consiste en la SEQ ID NO:1 o la SEQ ID NO:9, o comprendiendo dicha secuencia un fragmento del gen *mpdB* de 591 pares de bases que tiene la secuencia SEQ ID NO: 14.

13. Cepa bacteriana *Mycobacterium austroafricanum* registrada en la colección CNCM bajo el número I-3401.
14. Utilización del polipéptido de la SEQ ID NO:2 según la reivindicación 1, para la deshidrogenación del hidroxiisobutiraldehído a HIBA.
- 5 15. Procedimiento de tratamiento de efluentes acuosos que comprenden MTBE, ETBE y/o al menos uno de los intermedios catabólicos del MTBE, preferentemente seleccionado del grupo que consiste en el TBA, el 2-M-1,2-PD, el hidroxiisobutiraldehído y el HIBA, con el fin de reducir su concentración, caracterizado porque comprende la siembra de dicho efluente acuoso por al menos una célula según una cualquiera de las reivindicaciones 11 o 13 o identificada por el procedimiento según la reivindicación 12.

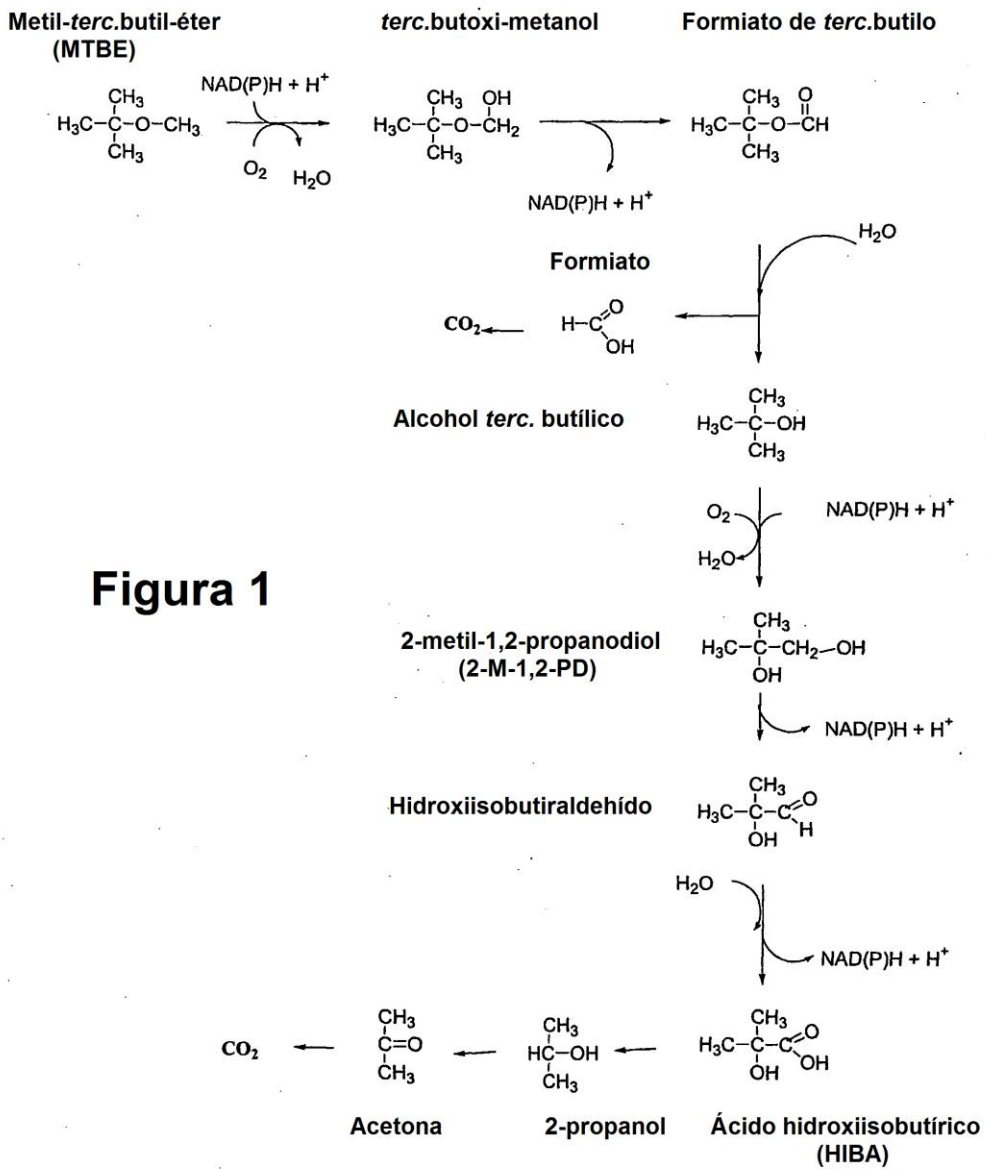


Figura 1

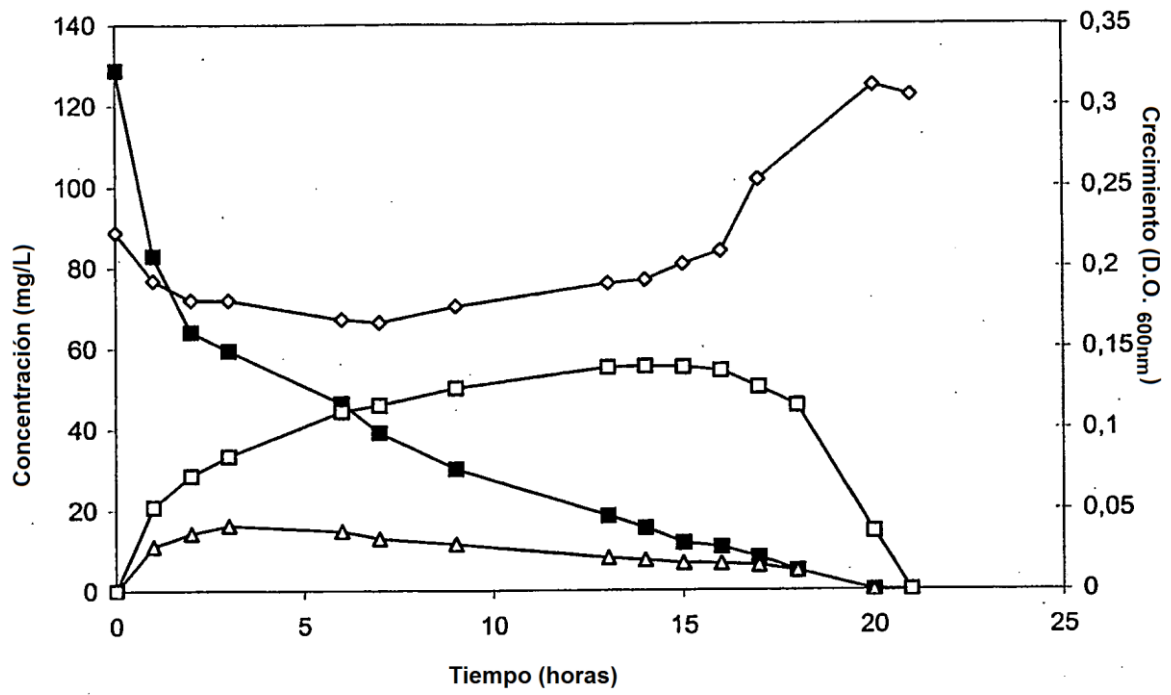
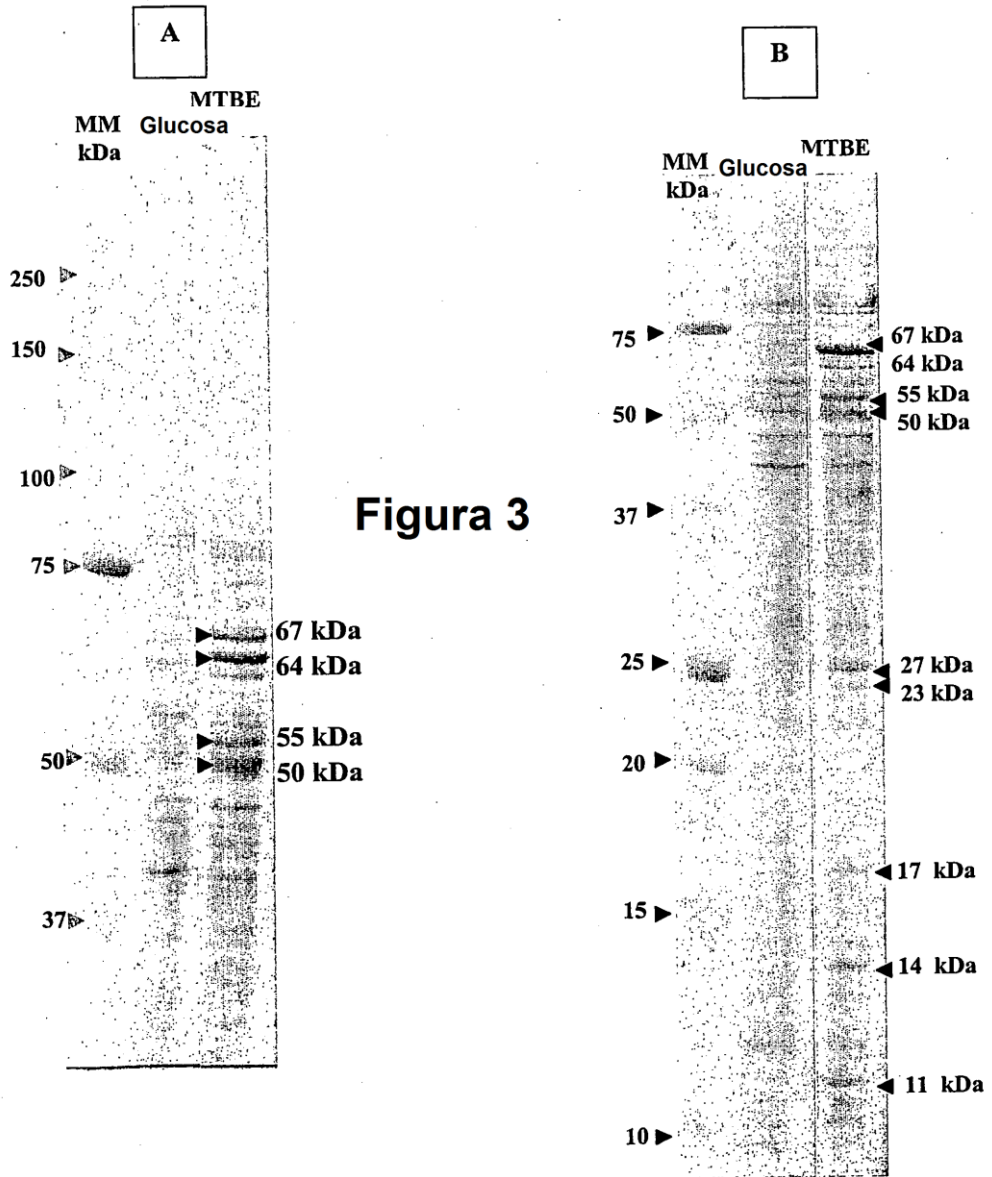


Figura 2



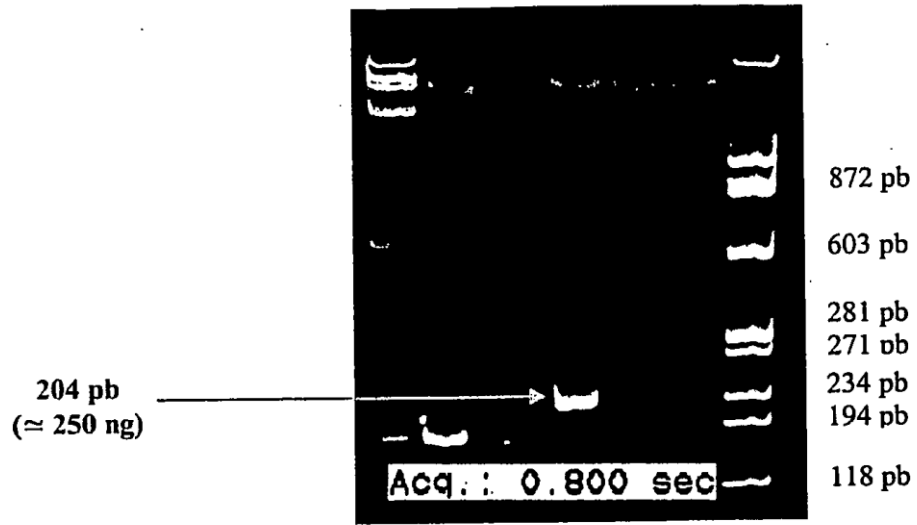


Figura 4

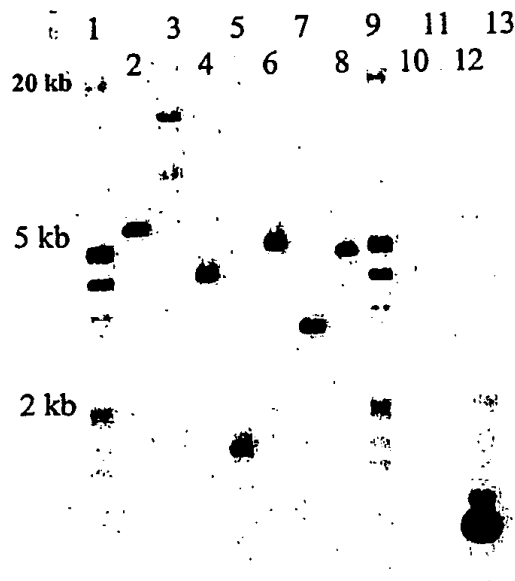


Figura 5

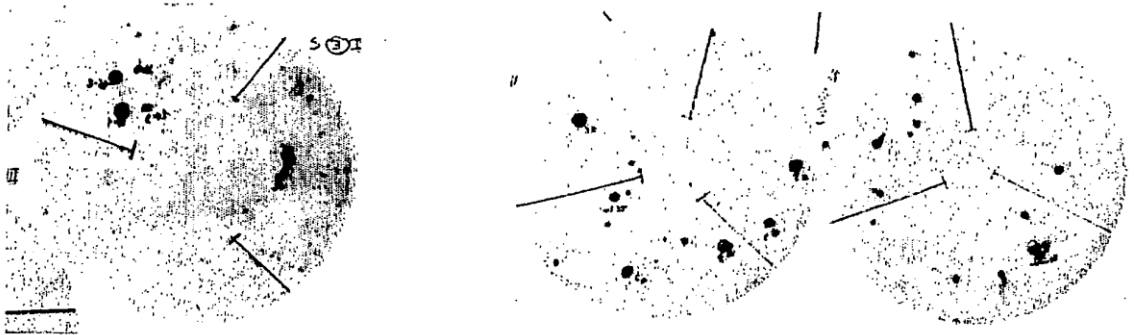
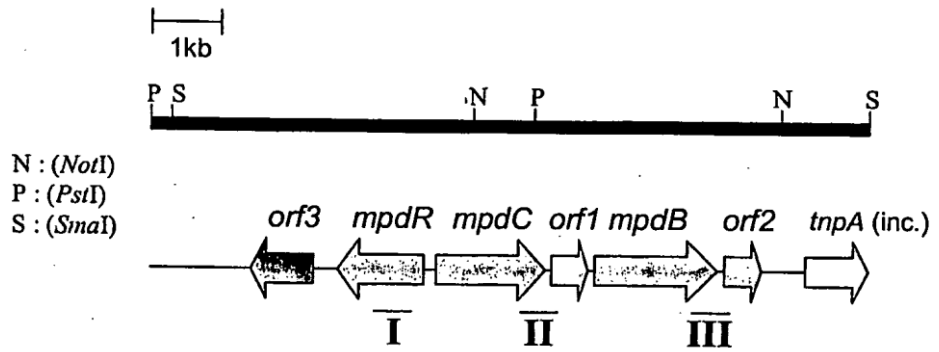


Figura 6



Grupo de genes *mpd*

Figura 7

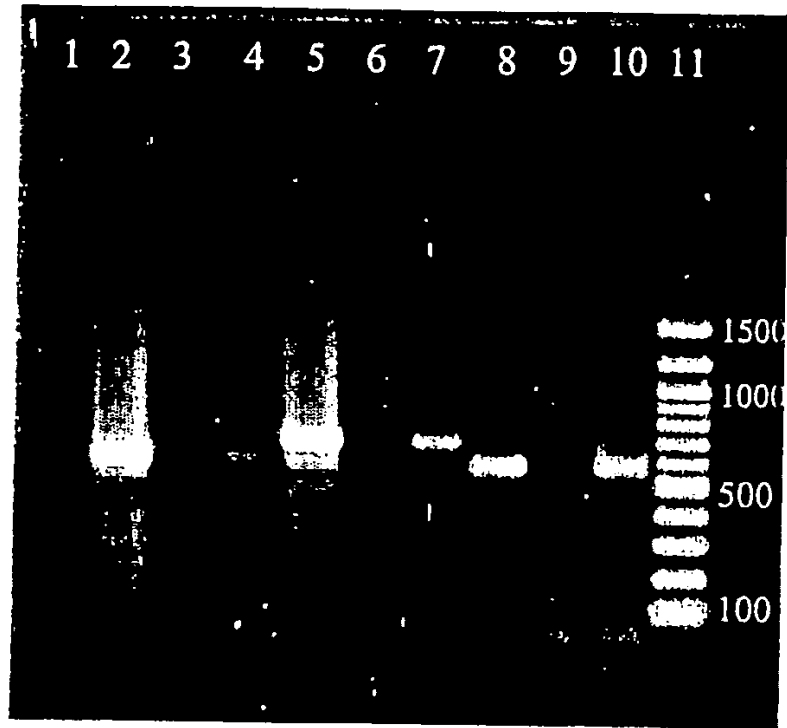


Figura 8

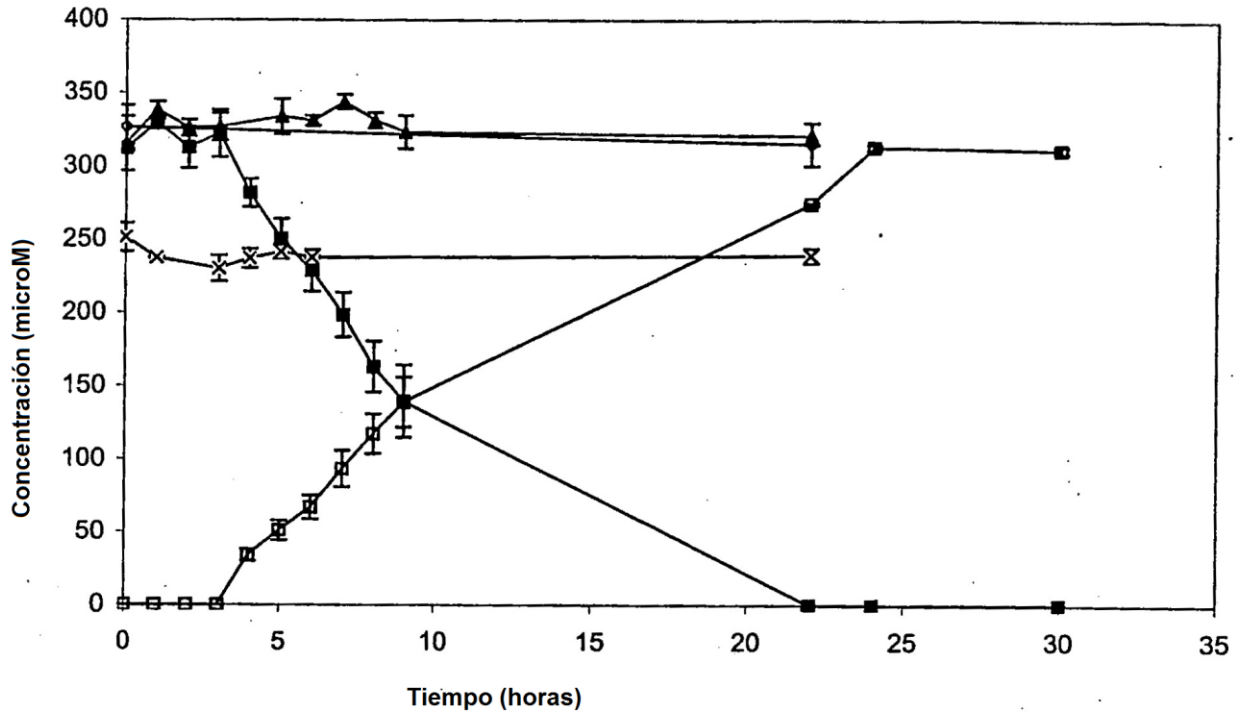


Figura 9

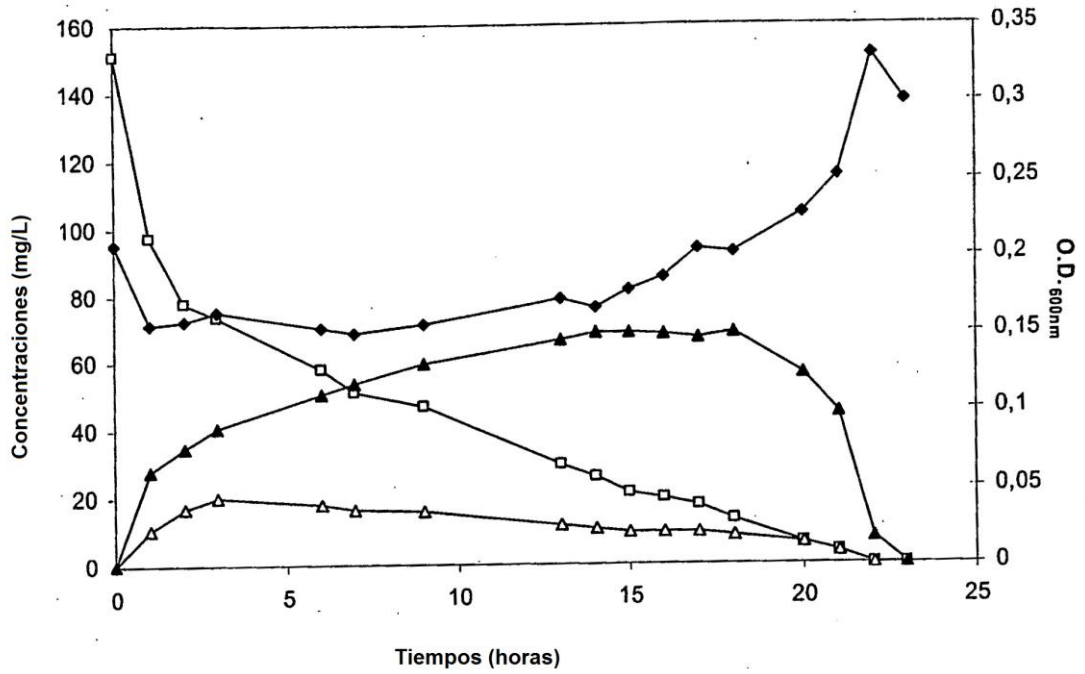


Figura 10

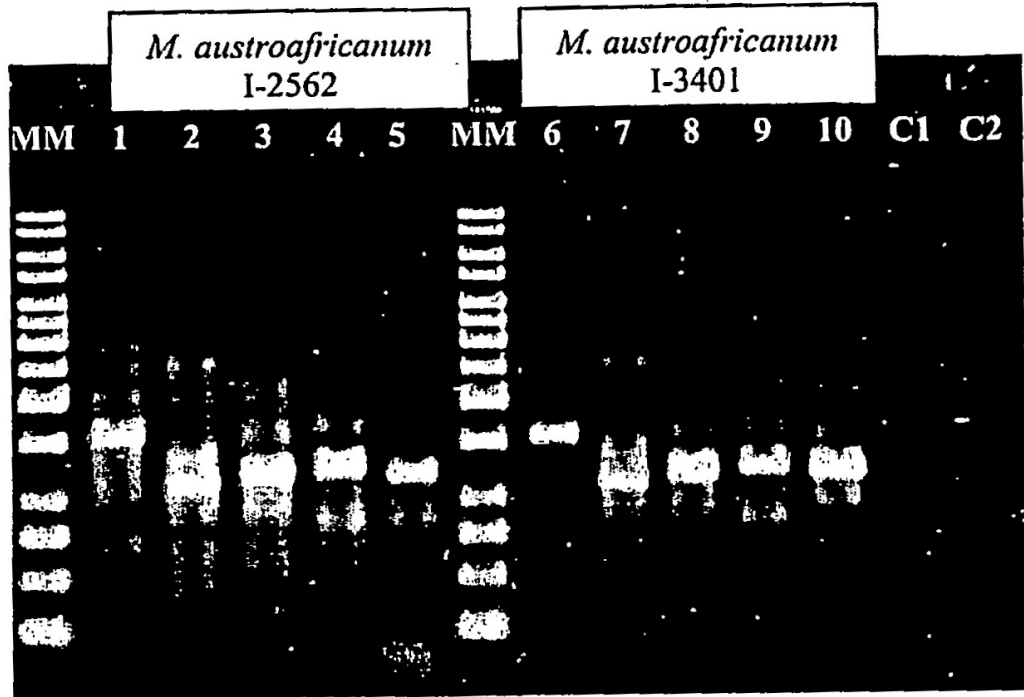


Figura 11

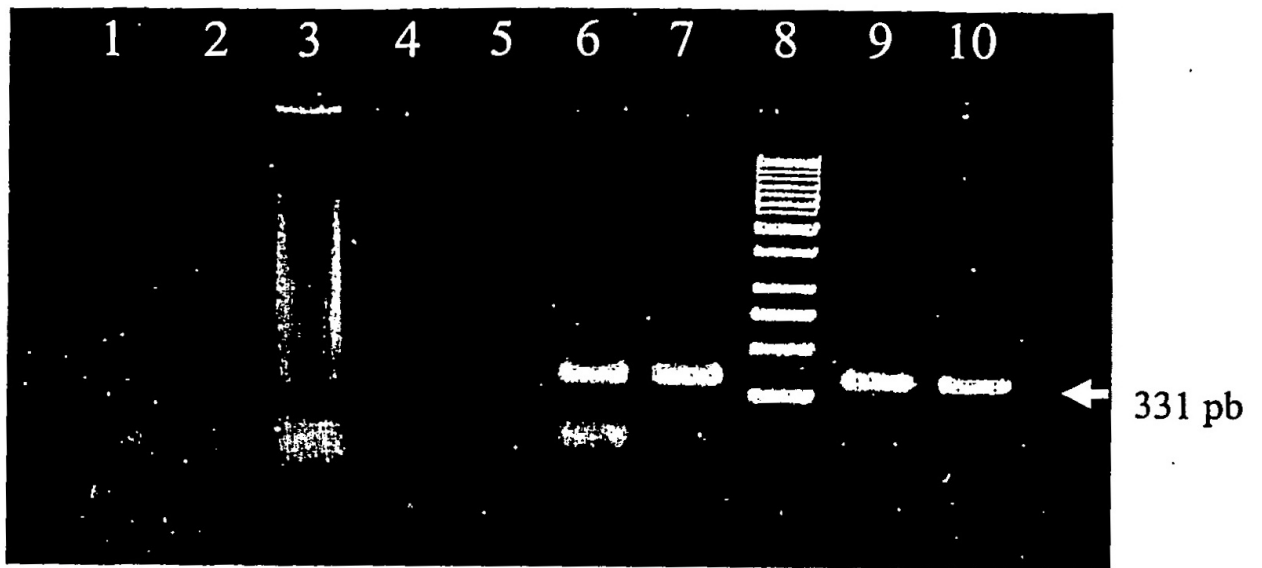


Figura 12