



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



① Número de publicación: **2 372 942**

② Número de solicitud: 201031068

⑤ Int. Cl.:
C07K 14/195 (2006.01)
C12N 9/90 (2006.01)
C12N 15/61 (2006.01)
C12P 23/00 (2006.01)

⑫

SOLICITUD DE PATENTE

A1

⑫ Fecha de presentación: **14.07.2010**

⑬ Fecha de publicación de la solicitud: **30.01.2012**

⑭ Fecha de publicación del folleto de la solicitud:
30.01.2012

⑰ Solicitante/s: **Universidad de Cantabria
Avda. de los Castros, s/n
39005 Santander, Cantabria, ES
Consejo Superior de Investigaciones Científicas,
Universitat de Barcelona y
Universidad de Almería**

⑱ Inventor/es: **Sangari García, Félix Javier;
García Lobo, Juan María;
Rodríguez Concepción, Manuel;
Pérez Gil, Jordi y
Carretero Paulet, Lorenzo**

⑳ Agente: **Zea Checa, Bernabé**

⑳ Título: **Nueva enzima para la biosíntesis de isoprenoides.**

㉑ Resumen:

Nueva enzima para la biosíntesis de isoprenoides.
Se ha encontrado una nueva enzima con actividad 1-de-soxi-D-xilulosa 5-fosfato reductoisomerasa que cataliza la reacción de producción de 2-C-metil-D-eritritol 4-fosfato a partir de 1-desoxi-D-xilulosa 5-fosfato, que tiene la secuencia de aminoácidos SEQ ID NO: 41. La enzima es útil en la síntesis de isoprenoides, particularmente en la síntesis de 2-C-metil-D-eritritol 4-fosfato.

DESCRIPCIÓN

Nueva enzima para la biosíntesis de isoprenoides.

5 La presente invención se enmarca de manera general en el campo de la biología molecular y la microbiología. En particular, la invención se refiere a enzimas involucradas en la síntesis de isoprenoides en bacterias.

Estado de la técnica

10 Los isoprenoides o terpenoides, uno de los grupos más abundantes de compuestos naturales, tienen papeles variados en la respiración, fotosíntesis, estructura de membranas, interacciones aleloquímicas, y regulación del crecimiento, entre otras. Todos los organismos de vida libre sintetizan isoprenoides a partir de los precursores de cinco átomos de carbono isopentenil difosfato (IPP), y su isómero en el doble enlace dimetilalil difosfato (DMAPP). Durante décadas se creyó que el IPP se sintetizaba exclusivamente a partir de acetil coenzima A mediante la ruta del mevalonato (MVA),
15 y que luego se convertía en DMAPP mediante una IPP/DMAPP isomerasa (IDI). Sin embargo, a principios de los años noventa del siglo pasado se descubrió que tanto el IPP como el DMAPP podían ser sintetizados simultáneamente a partir de piruvato y gliceraldehído 3-fosfato mediante una ruta alternativa conocida actualmente como ruta del 2-C-metil-D-eritritol 4-fosfato (MEP), o también como la ruta no-mevalonato.

20 En estos momentos está bien establecido que la mayoría de organismos emplean tan sólo uno de las dos rutas de síntesis de isoprenoides. De este modo, arqueas (Archaeobacteria), hongos y animales sintetizan IPP a partir de MVA, mientras que la mayoría de las bacterias (Eubacteria) sólo emplean la ruta MEP para la producción de precursores isoprenoides. Las plantas emplean ambas rutas, pero en diferentes compartimentos celulares: la ruta MVA sintetiza precursores isoprenoides citosólicos, mientras que la ruta MEP está localizada en los plástos.

25 Dado que la ruta MEP está ausente en los animales (incluidos los humanos) pero es esencial en un gran número de importantes patógenos bacterianos, se ha propuesto como una nueva diana prometedor para el desarrollo de nuevos agentes anti-infecciosos. Sin embargo, la información acerca de los posibles mecanismos de resistencia a un bloqueo de la ruta MEP es muy escasa. La resistencia a antibióticos puede ser originada por una exportación activa o un bloqueo en la entrada del fármaco, por su inactivación dentro de la célula, por modificación genética de su diana proteica, o por el uso de una ruta alternativa no afectada por el inhibidor, por mencionar tan solo algunas posibilidades. El inhibidor mejor caracterizado de la ruta MEP es la fosmidomicina (FSM), identificada inicialmente como un antibiótico natural eficaz contra un amplio espectro de bacterias. La FSM es un inhibidor específico de la 1-desoxi-D-xilulosa 5-fosfato (DXP) reductoisomerasa (DXR), la enzima que cataliza la producción de MEP a partir de DXP de un modo
30 dependiente de NADPH, en lo que constituye el primer paso específico de la ruta. La entrada de FSM en las células bacterianas es un proceso de transporte activo llevado a cabo por el transportador de glicerol 3-fosfato (GlpT) de un modo dependiente de cAMP. Un gen *glpT* deficiente en mutantes de *Escherichia coli*. o la ausencia de un homólogo de GlpT en otras bacterias como *Mycobacterium tuberculosis* conduce a una resistencia a FSM. La sobreexpresión del gen *fsr* de *E. coli*. que codifica una proteína similar a las proteínas bacterianas de exportación de fármacos, también conduce a la resistencia a FSM, probablemente porque esta proteína facilita la exportación del inhibidor. Más aún, se ha visto que varias mutaciones independientes son capaces de rescatar la supervivencia de cepas de *E. coli* deficientes en los dos primeros enzimas de la ruta MEP, DXP sintasa (DXS) y DXR, lo que sugiere que la bacteria puede responder a un bloqueo de estas actividades mediante el uso de otras enzimas que producen DXP o MEP cuando sufren mutaciones.

45 Además del estudio de las enzimas de la ruta MEP como dianas para el desarrollo de agentes anti-infecciosos, es igualmente de interés estudiar estas enzimas para mejorar los procedimientos a nivel industrial para la síntesis de isoprenoides. Los isoprenoides, gracias a su amplia diversidad estructural, poseen muchas aplicaciones en la industria, como por ejemplo, como fármacos, diluyentes, aromatizantes, biocombustibles, o insecticidas naturales. Algunos isoprenoides utilizados en la industria son los aceites esenciales, los carotenoides, los tocoferoles, el taxol, y la artemisina. Una alternativa muy prometedor para su producción industrial es la ingeniería metabólica de bacterias y plantas para ser utilizadas como biofactorías de isoprenoides de interés.

55 Hay un conocimiento muy limitado de las enzimas involucradas en la síntesis de isoprenoides en las diferentes especies (bacterias, plantas, etc). Por lo tanto, es deseable ampliar este conocimiento y proporcionar nuevas herramientas para avanzar en la síntesis y la aplicación industrial de estos compuestos.

Explicación de la invención

60 Los inventores han encontrado sorprendentemente una nueva clase de enzimas oxidoreductasas que cataliza la conversión de DXP a MEP para la síntesis de isoprenoides a través de la ruta MEP en células procariotas.

Los inventores han detectado que los genomas completamente secuenciados de un número de bacterias, incluida la patógena *Brucella abortus* 2308. contienen los genes de la vía MEP con la única excepción de la que codifica para DXR. La presente invención se refiere pues a la clonación del gen que codifica la enzima que es utilizada en estos organismos para producir MEP, la demostración de su actividad bioquímica tanto *in vivo* como *in vitro*. y la determinación de su distribución filogenética.

ES 2 372 942 A1

En esta descripción se utilizará DRL para denominar a la nueva clase de enzima identificada. DRL proviene de DXR-Like.

5 La secuencia de aminoácidos de la proteína DRL de *B. abortus* 2308 clonada coincide con la secuencia descrita en el NCBI con referencia Swiss-Prot Q2YIM3, que corresponde a la secuencia de aminoácidos predicha a partir de la secuencia de nucleótidos *BAB2 0264* con GeneID 3827542. En la referencia Q2YIM3 se describe esta secuencia como “oxidoreductasa putativa”. Es decir, la definición de la posible función de la secuencia de aminoácidos se realizó automáticamente a partir de una predicción de función comparando secuencias de bases de datos. De la misma manera, en la referencia *BAB2 0264* se describen como funciones putativas “homoserina deshidrogenasa predicha” y “proteína de unión a NAD(P)(+) con un dominio Rossmann-fold”.

15 Lo anterior implica que la secuencia de nucleótidos y aminoácidos como tal está descrita, pero la actividad propuesta para la proteína es el resultado de una predicción; es decir, no se ha clonado físicamente y se desconoce su auténtica función biológica.

20 Contrariamente, los inventores han clonado la secuencia de nucleótidos y han encontrado que la proteína codificada (DRL) es una proteína funcional. Los inventores han encontrado que DRL es una enzima con actividad 1-desoxi D-xilulosa 5-fosfato reductoisomerasa que cataliza la reacción de producción de MEP a partir de DXP de forma similar a la descrita para el enzima DXR.

25 La fosmidomicina, un inhibidor competitivo específico de DXR, inhibió el crecimiento de células de *B. abortus* que expresaban el transportador GltT de *E. coli* (requerido para la entrada de fosmidomicina), confirmando que existe en estas bacterias una actividad similar a DXR *in vivo* (DRL). Se encontró que la proteína DRL de *B. abortus* pertenece a una familia de proteínas no caracterizadas y similares en secuencia a la homoserina deshidrogenasa. Experimentos posteriores confirmaron que DRL y DXR catalizan la misma reacción bioquímica *in vitro*.

30 La enzima activa DRL de *B. abortus* se caracteriza por ser un homodímero con un peso molecular determinado por cromatografía de exclusión molecular de 80 kDa (aproximadamente el doble que el tamaño deducido de la secuencia proteica). DRL presenta una actividad máxima en un pH entre 7.5 y 8, y una temperatura óptima entre 40 y 45°C. La curva de parámetros enzimáticos de Lineweaver-Burk indica una $V_{\max} = 0.083 \mu\text{mol NADPH min}^{-1} \text{mg}^{-1}$, una $k_{\text{cat}} = 0.065 \text{ s}^{-1}$, y una $K_{\text{m(DXP)}} = 109 \mu\text{M}$ para la enzima recombinante DRL.

35 Se realizó un análisis filogenético y funcional entre diferentes especies de bacterias y se detectó que otras bacterias además de *B. abortus* poseen proteínas homologas a DRL, que también complementan funcionalmente a la cepa EcAB4-10 de *E. coli* deficiente en DXR. Estas enzimas se agrupan dentro del mismo clado filogenético según se describe a continuación.

40 Como se describe detalladamente más adelante en el apartado de realizaciones particulares, las búsquedas con BLAST de posibles secuencias homologas a DRL se llevaron a cabo en la base de datos UniProt. Sólo se consideraron significativos aquellos resultados que correspondían a valores de $E < 10^{-3}$.

45 Los análisis filogenéticos se llevaron a cabo a partir de alineamientos de secuencias de proteínas obtenidos con CLUSTALW con los métodos de Máxima Verosimilitud (Maximum Likelihood, ML), Neighbor-Joining (NJ) y Máxima Parsimonia (Maximum Parsimony, MP). Los análisis por ML fueron llevados a cabo en PHYML v2.4.5, usando el modelo JTT de evolución de proteínas. Los análisis por NJ y MP se implementaron en MEGA 4.0. usando los parámetros por defecto.

50 Posteriormente se comprobó que diez de las secuencias identificadas procedentes de diferentes cepas bacterianas encontradas a partir de los análisis tenían actividad DRL (según se deduce de su capacidad para complementar la cepa mutante de *E. coli* EcAB4-10).

55 Además, por otra parte, se realizó un alineamiento de estas diez secuencias aminoacídicas para generar un perfil de HMM usando el programa HMMER (<http://HMMER.janelia.org/>). El perfil de HMM fue usado para modelar las DRL y en búsquedas de posibles homólogos de DRL en grandes bases de datos de proteínas como Pfam o UNIPROT. El protocolo completo aparece detallado más adelante. El análisis a partir del perfil HMM es la forma más sensible y específica para identificar con seguridad a los homólogos funcionales de DRL.

60 Así, la invención proporciona una enzima con actividad 1-desoxi D-xilulosa 5-fosfato reductoisomerasa que cataliza la reacción de producción de MEP a partir de DXP, que tiene la secuencia de aminoácidos SEQ ID NO: 41.

65 Para saber si una secuencia aminoacídica nueva (“secuencia incógnita”) corresponde a una DRL funcional o bien si se quieren buscar en bases de datos nuevas enzimas DRL no identificadas hasta ahora, primero se realiza un BLASTp a partir de una secuencia query seleccionada del grupo que consiste en SEQ ID NO: 41-50 y los parámetros para el BLAST indicados a continuación en una base de datos de proteínas. Si se obtiene un valor de $E < 10^{-3}$, la secuencia incógnita sería homologa a las DRL de SEQ ID NO: 41-50.

ES 2 372 942 A1

En particular, (a) se lanza un BLASTp contra una base de datos de proteínas con los parámetros por defecto: umbral esperado = 10 (en inglés “expect threshold”), tamaño de palabra = 3 (“word size”), matriz = BLOSUM62 (“matrix”) coste de hueco = existencia:11 extensión:1 (“gap cost = existence:11 extension:1”), y utilizando como secuencia pregunta (“query”) una de las secuencias SEQ ID NO: 41-50; y se seleccionan las secuencias obtenidas que tienen un valor esperado $E < 10^{-3}$; y (b) se hace un análisis filogenético de las secuencias obtenidas en el paso (a) mediante alineamiento de las secuencias con CLUSTALW con los métodos de Máxima Verosimilitud, Neighbor-Joining y/o Máxima Parsimonia; y se seleccionan las secuencias que pertenecen al mismo clado filogenético al que pertenecen las secuencias SEQ ID NO: 41-50.

Como alternativa, puede comprobarse si la secuencia incógnita se ajusta al perfil de HMM con el programa “hmm-search” del paquete de aplicaciones HMMER, v. 2.3.3. Si se obtiene un valor $E < 10^{-3}$, se ajusta al perfil.

El siguiente paso es comprobar si esta secuencia agrupa dentro del clado filogenético DRL, que es el que agrupa a las DRL funcionales (es decir, aquellas capaces de sintetizar MEP a partir de DXP). Para ello se hace un análisis filogenético de las secuencias obtenidas a partir del BLAST o del perfil de HMM, mediante alineamiento de las secuencias con CLUSTALW con los métodos de Máxima Verosimilitud, Neighbor-Joining y/o Máxima Parsimonia; y se seleccionan las secuencias que pertenecen al mismo clado filogenético al que pertenecen las secuencias SEQ ID NO: 41-50.

En una realización particular, adicionalmente, se confirma la función *in vivo* de la secuencia incógnita identificada expresándola en una cepa defectiva en DXR (como por ejemplo la cepa de *E. coli* EcAB4-10). Si es capaz de complementar la pérdida de función de DXR en la cepa mutante, la enzima con secuencia incógnita es una DRL. Alternativamente, se puede confirmar la función de la secuencia incógnita identificada *in vitro* utilizando la correspondiente proteína purificada para ensayos de actividad enzimática en presencia de DXP, $MgCl_2$, DTT y NADPH. Se seleccionan aquellas secuencias que son capaces de oxidar el NADPH y producir MEP en estas condiciones. Estos procedimientos se describen en detalle más adelante en el apartado de realizaciones particulares.

En una realización de la invención, las enzimas, cuyas secuencias están descritas actualmente en las bases de datos de proteína, que cumplen con las características antes indicadas y que por lo tanto son DRL funcionales, son las que tienen una secuencia seleccionada del grupo que consiste en SEQ ID NO: 41-50.

En otra realización particular, la enzima tiene la secuencia SEQ ID NO: 41.

En otra realización particular, la enzima DRL es de *B. abortus*. y más particularmente de la cepa *B. abortus* 2308. En otra realización particular, la enzima activa es un homodímero que tiene un peso molecular de 80 kDa y una actividad óptima a un pH de entre 7.5-8 y a una temperatura de entre 40-45°C.

Otro aspecto de la invención se refiere a un vector de expresión que comprende la secuencia de DNA que codifica para las enzimas DRL antes explicadas y que permite la expresión de la enzima DRL funcionalmente activa. Particularmente, la secuencia de DNA que codifica la enzima DRL es la secuencia del gen *BAB2 0264* (SEQ ID NO: 51). El vector de expresión es normalmente un plásmido o un cassette que se inserta en una célula hospedadora para la expresión de la enzima. Para permitir la expresión de la enzima DRL, la secuencia de DNA que la codifica se une operativamente a una secuencia promotora que es capaz de dirigir la expresión en la célula hospedadora deseada. El término “unido operativamente” se refiere a la asociación de secuencias de ácido nucleico en un único fragmento de ácido nucleico de manera que la función de una secuencia es afectada por la otra. Un promotor está unido operativamente a una secuencia codificante cuando es capaz de efectuar la expresión de esa secuencia codificante. Además, el vector puede comprender otros elementos genéticos para conseguir con éxito, la transformación, selección y propagación de las células hospedadoras que contienen la secuencia de interés. En una realización preferida el vector es el plásmido pET23b y la célula hospedadora es *E. coli* BL21 (DE3)pLys.

La enzima DRL se obtiene por un método que comprende (a) transformar establemente células hospedadoras con la construcción de expresión que comprende la secuencia que codifica para la enzima DRL; y (b) cultivar las células en condiciones que permitan la expresión de la enzima. A partir de aquí, puede utilizarse el cultivo para la producción de isoprenoides.

Opcionalmente, a partir del cultivo celular puede aislarse y purificar la enzima para ser utilizada posteriormente para procedimientos industriales de síntesis de isoprenoides basados en la disponibilidad de una fuente abundante de enzima purificada. Al contrario que DXR, cuya acumulación está altamente regulada, la no existencia de homólogos de DRL en muchas eubacterias y en plantas permite su sobreproducción industrial a niveles elevados.

Otro aspecto de la invención se refiere al uso de la enzima DRL para la síntesis de isoprenoides y particularmente para la síntesis de MEP.

Así, otro aspecto se relaciona con un método de obtención de MEP que comprende: (a) expresar la enzima activa en un sistema celular; y (b) cultivar el sistema celular del paso (a) bajo condiciones y en un medio que permiten la producción y acumulación de MEP; por ejemplo, en cepas deficientes en actividad MEP citidililtransferasa que son incapaces de metabolizar el MEP.

También puede obtenerse MEP (a) proporcionando un sistema celular que permita la expresión de la enzima; (b) cultivando el sistema celular del paso (a) bajo condiciones que permiten la producción de la enzima; y (c) cultivando el sistema celular bajo condiciones y en un medio que permiten la producción y acumulación de MEP o bien purificando el enzima y utilizarlo en un sistema *in vitro* bajo condiciones y en un medio que permiten la obtención de MEP, donde el medio comprende un sustrato seleccionado del grupo que consiste en DXP y una mezcla de piruvato y gliceraldehído 3-fosfato. Si se utiliza la mezcla de piruvato y gliceraldehído 3-fosfato, el sistema *in vitro* comprende además enzima DXS purificada que permita el primer paso a DXP.

Otro aspecto de la presente invención es utilizar la nueva enzima DRL como target para identificar/diseñar inhibidores que puedan ser útiles como antibióticos contra diversos microorganismos. Así, la invención se relaciona también con un método de cribado para identificar potenciales agentes inhibidores de la enzima DRL, que comprende poner en contacto DRL con el potencial agente y analizar si hay inhibición. Por ejemplo, un método para evaluar si un agente tiene la capacidad de inhibir la enzima, comprende los pasos de (a) obtener la enzima purificada o no; (b) tratar la enzima DRL con el agente potencial inhibidor; (c) comparar la actividad de la DRL tratada con la actividad de una DRL no tratada, seleccionando los compuestos que tienen actividad inhibitoria. DRL lleva a cabo la conversión de DXP en MEP con consumo concomitante de NADPH. La caída de la concentración de NADPH refleja el nivel de actividad de DRL, y puede ser seguida fácilmente mediante métodos ópticos, midiendo la absorbancia a 340 nm. Este método podría ser fácilmente escalable para realizar un cribado de alto rendimiento (“high-throughput screening”, HTS).

A lo largo de la descripción y las reivindicaciones la palabra “comprende” y sus variantes no pretenden excluir otras características técnicas, aditivos, componentes o pasos. Para los expertos en la materia, otros objetos, ventajas y características de la invención se desprenderán en parte de la descripción y en parte de la práctica de la invención. Las siguientes realizaciones preferidas se proporcionan a modo de ilustración, y no se pretende que sean limitativas de la presente invención. Además, la presente invención cubre todas las posibles combinaciones de realizaciones particulares y preferidas aquí indicadas.

Descripción de los dibujos

La Fig. 1 muestra un esquema de la ruta MEP en bacterias. DXP, 1-desoxi-D-xilulosa 5-fosfato; MEP, 2-C-metil-D-eritritol 4-fosfato; CDP-ME, citidina difosfometilerytritol; IPP, isopentenil difosfato; DMAPP, dimetilalil difosfato; Pir, piruvato; G3P, gliceraldehído 3-fosfato; B₆/B₁₂, vitaminas B₆/B₁₂; Is, isoprenoides. Las enzimas están indicadas en negrita: DXS, DXP sintasa (EC 2.2.1.7); DXR, DXP reductoisomerasa (EC 1.1.1.267); MCT, MEP citidiltransferasa (EC 2.7.7.60). El paso inhibido por fosmidomicina (FSM) está señalado. Las flechas interrumpidas cerradas representan varios pasos. Las cepas de *E. coli* deficientes en las actividades DXS (EcAB4-2), DXR (EcAB4-10) o MCT (EcAB4-7) usadas aquí fueron construidas para utilizar MVA aportado exógenamente para la síntesis de IPP (representado con una flecha interrumpida abierta). La flecha puntuada cerrada marca la reacción catalizada por la enzima “DXR-like” (DRL).

La Fig. 2 muestra la complementación de cepas de *E. coli* deficientes en actividad DXR (EcAB4-10), DXS (EcAB4-2) o MCT (EcAB4-7) con un plásmido que expresa el gen *BAB2_0264* (pET-DRL). Los experimentos control fueron hechos con plásmidos que expresan los genes de *E. coli* que codifican DXR, DXS o MCT, y también los vectores vacíos (Ø). La capacidad del gen clonado para rescatar el crecimiento de las correspondientes cepas mutantes (es decir, su capacidad de complementación) fue establecida mediante la siembra por estría de colonias individuales en placas de cultivo suplementadas (+) o no (-) con 1 mM MVA como aparece indicado.

La Fig. 3 (A) muestra los cromatogramas de seguimiento de reacciones múltiples (“Multiple-reaction monitoring”, MRM) de una mezcla de reacción que contiene MgCl₂, DTT, NADPH, y DXP antes (0 h) y 2 h después de añadir DRL recombinante. El panel inferior muestra el resultado después de la incubación durante 2h en ausencia de enzima, o con DXR purificada en lugar de DRL. El pico 1 corresponde a DXP (m/z 213/97, tiempo de retención ca. 9.15 min). El pico 2 corresponde a MEP (m/z 215/97, tiempo de retención ca. 9.90 min), como se demuestra comparando su espectro de masas con el del patrón de MEP (B). “a MEP” significa “MEP auténtico”. (C) Cambios en los niveles de NADPH monitorizados por absorbancia a 340 nm en mezclas de reacción como las descritas en (B) después de añadir DRL recombinante (círculos). El inhibidor de DXR FSM se incorporó en las mezclas a la concentración indicada (μM). Se muestra también un control sin DXP (cuadrados blancos). “R NADPH L” es “niveles de NADPH relativos”. “P2” es “pico 2” y “In” es “Intensidad”.

La Fig. 4 muestra el análisis filogenético de los supuestos homólogos de DRL. Se construyó un árbol filogenético tras el análisis de Máxima Verosimilitud de las relaciones evolutivas entre los supuestos homólogos de DRL que se obtuvieron de las búsquedas BLAST con DRL de *B. abortus* como interrogante (Tabla 3). Los resultados están representados como un árbol circular sin raíz, dibujado a escala, con la longitud de las ramas proporcional al tiempo de evolución. Las secuencias que complementan la cepa de *E. coli* EcAB4-10 aparecen indicadas con un círculo negro, mientras que aquéllas que no complementan al mutante están indicadas con un círculo blanco. La letra dentro de los círculos indica si la secuencia pertenece a un organismo que carece (A) o no (B) de DXR. El clado DRL que agrupa a todas las DRLs activas aparece resaltado.

La Fig. 5 muestra la determinación del valor de la concentración mínima inhibitoria (“Minimum Inhibitory Concentration”, MIC) de FSM en *B. abortus* 2308. Diluciones consecutivas de FSM en medio BB fueron inoculadas con 10^5 bacterias/ml, e incubadas durante 24 h a 37°C. Se incluyó un tubo control sin FSM. En la fila superior se muestra el resultado obtenido con la cepa silvestre *B. abortus* 2308: en la fila inferior, la misma cepa transformada con el plásmido pFJS251 que porta el gen *glpT* de *E. coli*.

La Fig. 6 muestra el mapa de las regiones genómicas de *B. abortus* clonadas en los plásmidos señalados que rescataron el crecimiento de las células de *E. coli* deficientes en DXR. Las regiones codificantes están representadas por flechas. La flecha correspondiente al gen *BAB2 0264* (que codifica DRL) está coloreada en gris.

La Fig. 7 muestra (A) el análisis por SDS-PAGE de la sobreexpresión y purificación de la proteína DRL de *B. abortus* fusionada a una cola de polihistidinas. Las calles corresponden a extractos celulares de *E. coli* BL21(DE3)pLys llevando la construcción pET-DRL antes (-) y después (+) de la inducción con IPTG; P, fracción precipitada (insoluble); S, fracción sobrenadante (soluble); E, fracción eluída con 150 mM de imidazol después de una cromatografía de afinidad en columnas de níquel. La posición de los marcadores de peso molecular se indica a la izquierda. (B) Determinación del peso molecular de DRL de *B. abortus* mediante cromatografía de exclusión por tamaños. DRL fusionada a una cola de histidinas y producida en *E. coli* fue purificada, cargada en una columna de Superdex 200 HL 16/60, y eluída con 100 mM Tris-HCl pH=7.5, 300 mM NaCl, 1 mM DTT, y 2 mM MgCl₂ a 1 ml/min. El cromatograma muestra el pico de DRL que eluye a 76.5 min. (C) Gráfico de calibración usando los siguientes estándares: ribonucleasa A (13.7 kDa), anhidrasa carbónica (29 kDa), ovoalbúmina (44 kDa), conalbúmina (75 kDa), aldolasa (158 kDa), y ferritina (440 kDa). El peso molecular del pico de DRL se estimó ca. 80 kDa mediante análisis de regresión lineal de los tiempos de retención frente al peso molecular. “Ret t” es “tiempo de retención”.

La Fig. 8 muestra el cálculo de pH y temperatura óptimos de DRL de *B. abortus*. La actividad fue monitorizada en mezclas de reacción (100 μ l) que contienen 100 mM Tris, 1.5 mM MgCl₂, 1 mM DTT, 0.15 mM DXP, 0.2 mM NADPH y 40 μ g de enzima recombinante purificada, bien a 37°C en un rango de pH (A) o a diferentes temperaturas a un pH fijo de 7.5 (B). La actividad enzimática se dedujo de la disminución en la absorbancia a 340 nm conforme el NADPH fue oxidado, y representada de modo relativo al valor más alto. Se muestran la media y el error estándar de tres (n=3) ensayos. (C) Cálculo de las constantes de Michaelis-Menten. Los ensayos de actividad se llevaron a cabo a 40°C en 100 μ l de una mezcla de reacción que contiene 100 mM Tris pH 7.5, 1.5 mM MgCl₂, 1 mM DTT, 0.2 mM NADPH (preparado fresco) y variando las concentraciones de DXP (25-500 μ M) en presencia de 60 μ g de DRL de *B. abortus* recombinante purificada. La velocidad fue calculada como disminución en unidades de absorbancia a 340 nm por minuto en un rango lineal. La curva recíproca doble (Lineweaver-Burk) se hizo con los valores medios de experimentos triplicados (n=3).

La Fig. 9 muestra la complementación de células *E. coli* deficientes en DXR con las secuencias de DRL y DXR de bacterias de clase B. Las secuencias indicadas de DRL y DXR de *Roseobacter litoralis* Och 149, *Bacillus halodurans* C-125, y *Listeria monocytogenes* F2365 fueron amplificadas por PCR a partir de DNA genómico con los conjuntos de primers descritos en la Tabla 1 y clonados en pJET1.2. Las correspondientes construcciones y el vector vacío control (\emptyset) se utilizaron para transformar células EcAB4-10. La habilidad del gen clonado para rescatar el crecimiento de las cepas deficientes en DXR se averiguó siguiendo el crecimiento en placas suplementadas (+) o no (-) con 1mM MVA como se indica.

45 Exposición detallada de realizaciones particulares

Cepas bacterianas, medios y reactivos

La cepa *B. abortus* 2308 (NCBI taxonomy ID 359391) fue crecida en caldo Brucella (Brucella broth, BB) o placas de agar Brucella (Brucella agar, BA) (Pronadisa). Las cepas de *E. coli* fueron crecidas en caldo o placas de medio Luria-Bertani (LB). Cuando era requerido, los medios fueron suplementados con los siguientes antibióticos: 25 μ g/ml kanamicina, 100 μ g/ml ampicilina, 20 μ g/ml cloranfenicol. Cuando se indicó, el medio de cultivo fue también suplementado con diferentes concentraciones de FSM (Molecular Probes) o 1 mM MVA preparado a partir de un concentrado de mevalonato como está descrito (cf. M. Rodríguez-Concepción *et al.*, “Genetic evidence of branching in the isoprenoid pathway for the production of isopentenyl diphosphate and dimethylallyl diphosphate in *Escherichia coli*” *FEBS Lett* 2000. vol. 473, pp. 328-332). A menos que se indique de otra manera, los productos químicos y reactivos fueron obtenidos de Sigma-Aldrich. Las enzimas de restricción y de modificación de DNA fueron compradas de Promega. Los oligonucleótidos fueron sintetizados por Sigma-Aldrich, y aparecen mostrados en la Tabla 1.

ES 2 372 942 A1

TABLA 1

Oligonucleótidos utilizados

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

Organismo	Nombre	Secuencia
Escherichia coli DH5 α	GlpT_Sall.F	SEQ ID NO: 1
	GlpT_PstI.R	SEQ ID NO: 2
Brucella abortus 2308	DRL_NdeI.F	SEQ ID NO: 3
	DRL_XhoI.R	SEQ ID NO: 4
Ochrobactrum anthropi LMG3331	Oant_1387.DRL.F	SEQ ID NO: 5
	Oant_1387.DRL.R	SEQ ID NO: 6
Ochrobactrum anthropi LMG3331	Oant_4118.DRL.F	SEQ ID NO: 7
	Oant_4118.DRL.R	SEQ ID NO: 8
Mesorhizobium loti MAFF303099	MLL3630.DRL.F	SEQ ID NO: 9
	MLL3630.DRL.R	SEQ ID NO: 10
Mesorhizobium loti MAFF303099	MLR6497.DRL.F	SEQ ID NO: 11
	MLR6497.DRL.R	SEQ ID NO: 12
Bartonella henselae str Houston	BH12960.DRL.F	SEQ ID NO: 13
	BH12960.DRL.R	SEQ ID NO: 14
Roseobacter litoralis Och 149	RLO149_05338.DRL.F	SEQ ID NO: 15
	RLO149_05338.DRL.R	SEQ ID NO: 16
	RLO149_21444.DXR.R	SEQ ID NO: 17
	RLO149_21444.DXR.F	SEQ ID NO: 18
Bacillus halodurans C-125	BH0774.DRL.F	SEQ ID NO: 19
	BH0774.DRL.R	SEQ ID NO: 20
	BH2421.DXR.R	SEQ ID NO: 21
	BH2421.DXR.F	SEQ ID NO: 22
Listeria monocytogenes F2365	LMOF2365_0575.DRL.F	SEQ ID NO: 23
	LMOF2365_0575.DRL.R	SEQ ID NO: 24
	LMOF2365_1334.DXR.R	SEQ ID NO: 25
	LMOF2365_1334.DXR.F	SEQ ID NO: 26
Finegoldia magna ATCC 29328	FMG_0520.DRL.F	SEQ ID NO: 27
	FMG_0520.DRL.R	SEQ ID NO: 28
Nostoc punctiforme PCC 73102	Npun_R1071.DRL.R	SEQ ID NO: 29
	Npun_R1071.DRL.F	SEQ ID NO: 30
Anabaena variabilis PCC 7937	Ava_1114.DRL.F	SEQ ID NO: 31
	Ava_1114.DRL.R	SEQ ID NO: 32
Agrobacterium tumefaciens C58	Atu3161.DRL.F	SEQ ID NO: 33
	Atu3161.DRL.R	SEQ ID NO: 34
Candidatus Pelagibacter ubique HTCC1002	PU1002_02141.DRL.F	SEQ ID NO: 35
	PU1002_02141.DRL.R	SEQ ID NO: 36
Mycobacterium smegmatis mc(2)155	MSMEG_0894.F	SEQ ID NO: 37
	MSMEG_0894.R	SEQ ID NO: 38
Burkholderia cepacia J2315	BCAL1926.DRL.F	SEQ ID NO: 39
	BCAL1926.DRL.R	SEQ ID NO: 40

ES 2 372 942 A1

Clonado del gen *glpT*

El gen *glpT* fue aislado a partir de *E. coli* DH5 α mediante oligonucleótidos específicos del gen, y polimerasa Vent. Un fragmento de DNA del tamaño esperado fue obtenido y clonado en pJET.2 (Fermentas) y su secuencia nucleotídica determinada para descartar posibles mutaciones introducidas por la reacción de PCR. Para su expresión en células de *B. abortus*, la secuencia de *glpT* fue subclonada en pBBR1 MCS.

Construcción y cribado de una librería de *B. abortus* 2308

El DNA genómico fue extraído de células *B. abortus* 2308 conforme estaba descrito previamente (cf. FJ Sangari *et al.*, "Identification of *Brucella abortus* B19 vaccine strain by the detection of DNA polymorphism at the ery locus" *Vaccine* 1994, vol. 12, pp. 435-438). Tras la digestión parcial con *Sau3A*, se purificaron a partir del gel fragmentos de entre 3 y 6 kb, que se ligaron a pUC19 previamente digerido con *Bam*HI y defosforilado (Fermentas). Con la mezcla de ligación se transformó a *E. coli* DH5 α , y se plaqueó en medio LB con ampicilina a una densidad de 500 ufc por placa para amplificar la librería. Se extrajo plásmido a partir de las células raspadas de las placas y mezcladas. Para el cribado, 1 μ g de este DNA se electroporó en la cepa deficiente en DXR EcAB4-10, y se seleccionaron los transformantes en placas de LB suplementadas con cloranfenicol (para seleccionar la interrupción del gen *dxr*), kanamicina (para seleccionar la presencia del operón MVA) y ampicilina (para seleccionar la incorporación de los plásmidos de la librería). Los plásmidos aislados de transformantes que eran capaces de formar colonias fueron secuenciados para comprobar la identidad de los insertos.

Clonado de las secuencias DRL y ensayos de complementación

El DNA genómico de bacterias que contenían posibles secuencias DRL fue amplificado con los pares de oligonucleótidos descritos en la Tabla 1. Fragmentos de DNA del tamaño esperado se purificaron y se analizaron mediante digestión con enzimas de restricción para confirmar su identidad. Los fragmentos positivos fueron clonados en pJET1.2 (Fermentas). El DNA plasmídico de dos clones independientes de cada construcción se usó para transformar células *E. coli* EcAB4-10. Se consideró que los fragmentos clonados codificaban enzimas DRL funcionales cuando permitieron el crecimiento de las células EcAB4-10 en ausencia de MVA. Cuando se indicó, las cepas de *E. coli* EcAB4-2 y EcAB4-7 fueron también usadas en ensayos de complementación como se describió previamente (cf. S. Sauret-Güeto *et al.*, "A mutant pyruvate dehydrogenase E1 subunit allows survival of *Escherichia coli* strains defective in 1-deoxy-D-xylulose 5-phosphate synthase" *FEBS Lett* 2006, vol. 580, pp. 736-740).

Producción de proteína DRL recombinante de *B. abortus*

Los oligonucleótidos con SEQ ID NO: 3-4 (Tabla 1) fueron usados para amplificar *BAB2 0264* sin el codón de parada a partir de DNA genómico de *B. abortus* 2308. El fragmento amplificado se clonó en el vector de expresión pET23b (Novagen) después de digerirlo con *Nde*I y *Xho*I. Después de la transformación de células de *E. coli* BL21 (DE3)pLys con la construcción resultante (pET-DRL), la producción de una proteína DRL quimérica con una cola de seis residuos de histidina en el extremo C-terminal fue inducida mediante la adición de 0.4 mM IPTG a cultivos con una DO₆₀₀=0.5. Después de 14 h de crecimiento a 28°C, las células bacterianas fueron recogidas por centrifugación, y las células sedimentadas fueron resuspendidas en Tampón A (40 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM DTT) suplementado con 1 mg/ml de lisozima, 0.5 mM EDTA y una pastilla de cóctel completo de inhibidores de proteasas (Roche) por cada 10 ml de tampón. Las células resuspendidas se incubaron a 4°C durante 10 min y tras una breve sonicación (4 pulsos de 30 sec a 30 W) el lisado celular fue centrifugado a 19.000 xg durante 20 min. El sobrenadante se incubó durante 10 min a 4°C con 1/7 volúmenes de sulfato de protamina al 1% (p/v) en agua y centrifugado a 37.000 xg durante 45 min. Siempre en una cámara fría, el sobrenadante limpio fue incubado durante 2 h con 2 ml de agarosa Ni-NTA (Qiagen) y la mezcla se cargó en una columna de cromatografía vacía "poly-prep" (Bio-Rad). Después de añadir 10 ml de tampón A y 5 ml de tampón de lavado (tampón A con 25 mM imidazol), la proteína DRL recombinante se eluyó con alícuotas de 0.75 ml de tampón de elución (tampón A suplementado con 150 mM imidazol). Las fracciones que contenían DRL se juntaron y guardaron a -20°C en glicerol al 50%.

Caracterización bioquímica de la enzima DRL de *B. abortus*

La proteína recombinante DRL purificada se utilizó para realizar ensayos de actividad DXR en mezclas de reacción con 15 μ g de proteína purificada en Tris-HCl 100 mM pH 7.5, MgCl₂ 1,5 mM, DTT 1 mM, DXP 0.78 mM (Echelon) y NADPH 1 mM. Se prepararon reacciones control con 15 μ g de proteína DXR de *E. coli* recombinante o con agua. Tras incubación las reacciones a 37°C durante 2h, se diluyeron con agua en proporción 2:1 y la mezcla resultante se analizó por cromatografía líquida de altísimo rendimiento ("ultra-performance liquid chromatography", UPLC) acoplada a espectrometría de masas ("mass spectrometry", MS). La separación por UPLC se llevó a cabo en una columna Nucleodex β -OH 200 x 4 mm (Macherey Nagel) a un flujo de 0.75 ml/min con un divisor de flujo post-columna de 1:3. Se utilizó un equipo Acquity UPLC System (Waters) provisto de una bomba binaria con acetonitrilo como solvente A y acetato amónico 10 mM pH 6.5 como solvente B en un gradiente A:B de 9:1 a 4:6 de 0 a 15 min y un flujo isocrático de 9:1 hasta el minuto 20. Los análisis de MS y MS/MS se realizaron con un espectrómetro de masas de

triple cuadrupolo API 3000 (Applied Biosystems) usando la fuente de turbonebulización en modo de ión negativo con los siguientes valores: -3500 V de voltaje capilar, 8 unidades arbitrarias de gas nebulizador (N₂), 8 unidades arbitrarias de gas cortina (N₂), 4 unidades arbitrarias de gas de colisión (N₂), -30 V de potencial de desolvatación, potencial de enfoque de -200 V, potencial de entrada de -10 V, y energía de colisión de -30. Se realizó una adquisición de datos en modo de barrido completo desde m/z 50 a m/z 800 usando un tiempo de ciclo de 2 segundos, un tamaño de paso de m/z 0,1 y una pausa entre barridos de 5 milisegundos. Las moléculas DXP y MEP se identificaron en modo de reacción de monitorización múltiple (“multiple reaction monitoring”, MRM) usando las transiciones 213/97 para DXP y 215/97 para MEP en Q1 y Q3, respectivamente, con una resolución de unidad y una pausa entre barridos de 5 milisegundos en ambos casos. Para obtener los espectros de masa de cada pico, se llevó a cabo un barrido de iones hijos seleccionando la masa parental en Q1 y aplicando distintas energías de colisión para cada compuesto en Q3 (rampa lineal de -20 a -30 para DXP y de -20 a -35 para MEP).

Para el ensayo de inhibición con FSM y las estimaciones de los parámetros cinéticos, la actividad DXR se cuantificó a partir del cambio de absorbancia a 340 nm de la muestra de reacción como consecuencia de la oxidación del NADPH.

Análisis de la secuencia y análisis filogenético

Los genes codificantes de enzimas de biosíntesis de isoprenoides se obtuvieron mediante BLAST genómico a partir de la base de datos del NCBI. Las búsquedas con BLAST de posibles secuencias homologas a DRL se llevaron a cabo en la base de datos UniProt. Sólo se consideraron significativos aquellos resultados que correspondían a valores de $E < 10^{-3}$. Los análisis filogenéticos se llevaron a cabo a partir de alineamientos de secuencias de proteínas obtenidos con CLUSTALW (cf. JD Thompson *et al.*, “The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools”, *Nucleic Acids Res* 1997, vol. 25, pp. 4876-4882) con los métodos de Máxima Verosimilitud (Maximum Likelihood, ML), Neighbor-Joining (NJ) y Máxima Parsimonia (Maximum Parsimony, MP). Los análisis por ML fueron llevados a cabo en PHYML v2.4.5, usando el modelo JTT de evolución de proteínas (cf. D.T. Jones *et al.*, “The rapid generation of mutation data matrices from protein sequences” *Comput Appl Biosci* 1992, vol 8, pp. 275-282; and S. *et al.*, “A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood” *Syst Biol* 2003, vol. 52, pp. 696-704). Los análisis por NJ y MP se implementaron en MEGA 4.0. usando los parámetros por defecto (cf. K. Tamura *et al.*, “MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software versión 4.0.” *Mol Biol Evol* 2007, vol. 24, pp. 1596-1599). Para dotar de confianza estadística a la topología obtenida, se hizo un análisis de remuestreo (bootstrap analysis) con 1000 replicas en cada caso.

Células de *B. abortus* que expresan GlpT se vuelven sensibles a FSM

El genoma de *B. abortus* está completamente secuenciado, y carece de un gen codificante de DXR, pero sí que contiene un homólogo de DXS (la enzima que sintetiza DXP, el sustrato para DXR) así como genes codificantes de la MEP citidililtransferasa (MCT) y del resto de enzimas de la ruta MEP requeridas para transformar MEP en IPP y DMAPP (Fig. 1). Esto sugiere que una proteína de *B. abortus* que no presenta homología en conjunto a DXR podría ser la responsable de la transformación de DXP en MEP. Sin embargo, el crecimiento de células de *B. abortus* no fue inhibido por concentraciones de FSM de hasta 1 mg/ml. Este resultado sugiere que la enzima similar a DXR (“DXR-like”, DRL) podría no ser inhibida por FSM o, de modo alternativo, que el inhibidor era degradado, expulsado, o no incorporado por las células vivas. Consistente con esta última posibilidad, el genoma de *B. abortus* no contiene ningún homólogo al transportador GlpT, que ha sido implicado en la incorporación de FSM al interior celular. Para investigar si una incorporación deficiente de FSM era la causa del fenotipo de resistencia de esta bacteria, el gen codificante del transportador GlpT de *E. coli* fue expresado en células de *B. abortus*. Como se muestra en la Fig. 5, los transformantes se hicieron sensibles a FSM con una concentración mínima inhibitoria (MIC) de 4 μ g/ml, confirmando que el fenotipo resistente de las células silvestres era resultado de tan sólo la ausencia de un mecanismo de entrada apropiado. Estos datos también sugieren que la actividad de la posible proteína DRL de *B. abortus* era de hecho inhibida por FSM, consistente con la hipótesis de que el mecanismo bioquímico usado por esta enzima alternativa para producir MEP a partir de DXP pudiera ser similar al usado por DXR.

La complementación de una cepa de *E. coli* deficiente en DXR conduce a la identificación de DRL

Para identificar el gen codificante de DRL en *B. abortus*, se construyó una librería genómica de esta bacteria y se usó para complementar a un mutante de *E. coli* deficiente en DXR. El genoma de la cepa *E. coli* EcAB4-10 contiene una delección del gen *dxr* y un operón MVA sintético que permite la producción de IPP y DMAPP (y por lo tanto la supervivencia de las células) cuando se aporta MVA al medio de cultivo. Células competentes de EcAB4-10 se transformaron con la librería genómica de *B. abortus* y fueron plaqueadas en ausencia de MVA. Los plásmidos de transformantes positivos que crecieron sin MVA exógeno fueron secuenciados y mostraron contener fragmentos genómicos de *B. abortus* solapados que contenían los genes *BAB2 0264* y *BAB2 0265* (Fig. 6). *BAB2 0265* codifica para una hidrolasa de la familia HAD, mientras que *BAB2 0264* codifica para una proteína anotada como una oxidoreductasa putativa y fue seleccionado para experimentos posteriores. La transformación de células EcAB4-10 con un vector que contenía tan sólo *BAB2 0264* dio lugar a una complementación completa de la auxotrofia para MVA (Fig. 2), lo que sugiere que la proteína codificada (Q2YIM3) era la enzima DRL de *B. abortus* predicha. La misma construcción fue usada en experimentos de complementación con las cepas de *E. coli* EcAB4-2 y EcAB4-7, deficientes

en las actividades DXS y MCT respectivamente. Una proteína que catalice la misma reacción bioquímica que DXR no debería rescatar el crecimiento de las cepas EcAB4-2 (porque requeriría DXP, el producto de la actividad DXS) ni EcAB4-7 (ya que produciría MEP que necesita ser convertido en intermediarios posteriores de la ruta por la actividad MCT). Por el contrario, una proteína que usara un sustrato diferente de DXP para producir MEP o un intermediario posterior rescataría el crecimiento de la cepa EcAB4-2, deficiente en DXS, mientras que una proteína que produjera un intermediario de la ruta posterior a MEP a partir de DXP u otro sustrato rescataría el crecimiento de la cepa EcAB4-7, deficiente en MCT. Como se muestra en la Fig. 2, la expresión de *BAB2 0264* en estas cepas no rescata su auxotrofia para MVA. Estos resultados son una importante evidencia *in vivo* de que este gen codifica DRL, una enzima distinta de DXR que necesita el producto de DXS para sintetizar un producto requerido por la actividad MCT (i.e. usa DXP o un metabolito derivado para producir MEP o un precursor que podría ser transformado en MEP por *E. coli*).

DRL y DXR catalizan la misma reacción bioquímica

La secuencia identificada de la proteína DRL de *B. abortus* no muestra homología general con secuencias de DXR sino con enzimas del tipo homoserina deshidrogenasas (HD). Una búsqueda de dominios funcionales detectó la existencia de un dominio N-terminal de unión a NAD(P) con un plegamiento tipo Rossmann modificado, similar al que se encuentra en muchas deshidrogenasas. Esto, unido a los resultados *in vivo* (Fig 2 y Fig. 3A), sugería que DRL podría usar NADPH para catalizar una reacción muy similar (o incluso idéntica) a la catalizada por DXR. Para verificar esta posibilidad, se produjo en *E. coli* una versión recombinante de la proteína DRL de *B. abortus* fusionada a una cola de histidinas y se purificó (Fig. 3A) para ser usada en ensayos *in vitro* de actividad DXR. Ensayos control se realizaron con proteína recombinante purificada DXR de *E. coli* y los productos de reacción se identificaron por UPLC-MS(/MS). Se encontró que las muestras que contenían o DRL o DXR producían MEP a partir de DXP de forma similar (Fig. 3A). La identidad del producto de reacción en las muestras con DRL se confirmó comparando su patrón de fragmentación con el de un patrón de MEP (Fig. 3B). Cuando se prepararon mezclas de reacción sin enzima o sin NADPH solo se detectaron picos para DXP (Fig. 3A). Además, se observó que la actividad de la enzima DRL recombinante era sensible a la inhibición con FSM (Fig. 3C).

Experimentos de cromatografía de exclusión molecular mostraron que la enzima DRL activa es un homodímero (Fig. 7), como DXR (cf. PJ Proteau, “1-Deoxy-D-xylulose 5-phosphate reductoisomerase: an overview” *Bioorg Chem* 2004. vol. 32, pp. 483-493). También de forma similar a DXR, la actividad DRL tiene un pH óptimo de 7.5 a 8 y una temperatura óptima de 40 a 45°C (Fig. 8). La representación tipo Lineweaver-Burk para el cálculo de los parámetros cinéticos (Fig. 8) mostró una $V_{max} = 0.083 \mu\text{mol NADPH min}^{-1} \text{mg}^{-1}$, una $k_{cat} = 0.065 \text{ s}^{-1}$, y una $K_m(\text{DXP}) = 109 \mu\text{M}$ para la proteína DRL recombinante. En las mismas condiciones experimentales, para DXR de *E. coli* se observó una $K_m(\text{DXP}) = 211 \mu\text{M}$, del mismo orden que los valores reportados en la literatura para distintas enzimas DXR (CF. PJ Proteau, *supra*). En contraste, los valores de k_{cat} para DXR (CF. PJ Proteau, *supra*) son entre 2 y 3 órdenes de magnitud mayores que los calculados para la DRL recombinante de *B. abortus*. En conjunto, DRL define una nueva clase de enzimas dependientes de NADPH relacionadas con oxidoreductasas tipo HD que catalizan la formación de MEP a partir de DXP de forma casi idéntica a DXR, aunque la tasa de conversión parece ser menor en el caso de DRL.

La distribución de secuencias DRL y DXR no es mutuamente excluyente en todos los organismos

La búsqueda de secuencias de proteínas similares a la de DRL de *B. abortus* (Q2YIM3) en las bases de datos de UNIPROT con el algoritmo BLAST identificó un total de 185 secuencias (Tabla 3). La mayoría de las bacterias solo mostraron 1 secuencia con homología a DRL, pero 17 cepas mostraron 2 y 3 cepas mostraron 3 (Tabla 4). Cuando se comparó la distribución de secuencias DRL putativas con la de enzimas de la ruta del MEP, se establecieron 3 clases de organismos (Tabla 4).

La primera (clase A) estaba formada por bacterias con secuencias de DRL en lugar de DXR en sus genomas. La mayoría de estas secuencias DRL (incluyendo la proteína de *B. abortus* identificada) eran de alfa-proteobacterias, pero también se encontraron algunas en firmicutes. En particular, las secuencias únicas de DRL encontradas en los genomas de la alfa-proteobacteria *Bartonella henselae* (Q6G2D9) o la firmicute *Fingoldia magna* (B0S038) fueron activas en experimentos de complementación de la cepa EcAB4-10 (Tabla 2). El mismo abordaje confirmó la actividad de las secuencias duplicadas de DRL presentes en el genoma de las alfa-proteobacterias *Ochrobactrum anthropi* (A6WYQ0 y A6X6G6) y *Mesorhizobium loti* (Q98FT2 y Q989B6).

La segunda (clase B) estaba formada por bacterias con secuencias codificantes tanto para DRL como para DXR (Tabla 4). De estas, las secuencias DRL de las alfa-proteobacterias *Agrobacterium tumefaciens* (A9CES2) y *Candidatus Pelagibacter ubique* (Q1V2P9), la beta-proteobacteria *Burkholderia cepacia* (B4EB12), la actinobacteria *Mycobacterium smegmatis* (A0QQV9), y las cianobacterias *Nostoc punctiforme* (B21VC2) y *Anabaena variabilis* (Q3ME48) no rescataron la pérdida de actividad DXR en células de *E. coli* (Tabla 2). En contraste, las secuencias de la alfa-proteobacteria *Roseobacter litoralis* (A9HDV1) y las firmicutes *Listeria monocytogenes* (Q723A4) y *Bacillus halodurans* (Q9KES5) fueron activas en ensayos de complementación (Tabla 2), lo que sugería que estos organismos podrían tener enzimas redundantes catalizando la producción de MEP. Sin embargo, únicamente las secuencias homólogas a DXR de *L. monocytogenes* (Q720A5) y *B. halodurans* (Q9KA69) fueron capaces de complementar el mutante EcAB4-10 (Tabla 2 y Fig. 9). Por el contrario, la secuencia DXR de *R. litoralis* (A9GU34) era inactiva (Tabla 2 y Fig. 9), probablemente porque presenta muchos cambios de aminoácidos en posiciones que están altamente conservadas en

ES 2 372 942 A1

enzimas DXR funcionales de otras bacterias y plantas. Estos resultados indican que algunos organismos de clase B son funcionalmente equivalentes a los de la clase A en términos de la ruta MEP (i.e. tienen una enzima DRL activa pero carecen de una enzima DXR).

5 La tercera (clase C) estaba formada por bacterias con DRL pero sin enzimas de la ruta MEP. Este grupo incluye arqueobacterias y bacterias que no usan la ruta del MEP para la biosíntesis de isoprenoides (Tabla 4).

La Tabla 2 muestra los resultados de la complementación de la cepa EcAB4-10 de *E. coli*. que carece de DXR, con las secuencias indicadas. Los valores de "identidad" (%) se indican relativos a la proteínas DRL de *B. abortus* (Q2YIM3). La columna "clase" indica la presencia de homólogos de DRL, DXR y/o otras enzimas de la ruta del MEP en el mismo organismo: A (+DRL, -DXR, +MEP), B (+DRL, +DXR, +MEP), C (+DRL, -DXR, -MEP). La columna "C-DRL" indica si las correspondientes secuencias DRL complementan (+) o no (-) la cepa EcAB4-10. La columna "C-DXR" indica si las correspondientes secuencias DXR complementan (+) o no (-) la cepa EcAB4-10 (Fig. 9); np, secuencia DXR no presente en el genoma; nt, secuencia DXR no ensayada en experimentos de complementación. La secuencias están divididas entre las que pertenecen al clado DRL que se muestra en la Fig. 4, y las que no pertenecen al clado ("No clado").

A. N.	Identidad	Organismo (cepa)		Clase	C-DRL	C-DXR
Clado						
20	100	<i>Brucella abortus</i> 2308	SEQ ID NO: 41	A	+	np
	90	<i>Ochrobactrum anthropi</i> LMG3331	SEQ ID NO: 42	A	+	np
25	75	<i>Ochrobactrum anthropi</i> LMG3331	SEQ ID NO: 43	A	+	np
	73	<i>Mesorhizobium loti</i> MAFF303099	SEQ ID NO: 44	A	+	np
	67	<i>Mesorhizobium loti</i> MAFF303099	SEQ ID NO: 45	A	+	np
30	65	<i>Bartonella henselae</i> str Houston	SEQ ID NO: 46	A	+	np
	62	<i>Roseobacter litoralis</i> Och 149	SEQ ID NO: 47	B	+	-
35	43	<i>Bacillus halodurans</i> C-125	SEQ ID NO: 48	B	+	+
	40	<i>Listeria monocytogenes</i> F2365	SEQ ID NO: 49	B	+	+
40	33	<i>Fingoldia magna</i> ATCC 29328	SEQ ID NO: 50	A	+	np
NO Clado						
	36	<i>Nostoc punctiforme</i> PCC 73102		B	-	nt
45	35	<i>Anabaena variabilis</i> PCC 7937		B	-	nt
	35	<i>Agrobacterium tumefaciens</i> C58		B	-	nt
50	29	<i>Candidatus Pelagibacter ubique</i> HTCC1002		B	-	nt
	32	<i>Mycobacterium smegmatis</i> mc(2)155		B	-	nt
55	27	<i>Burkholderia cepacia</i> J2315		B	-	nt

55 *Los homólogos de DRL se agrupan en un grupo monofilogenético (clado)*

Cuando las secuencias aminoacídicas de DRL putativas identificadas en búsquedas con el algoritmo BLAST fueron sometidas a un análisis filogenético tipo "Maximum Likelihood" (ML), todas las secuencias de DRL activas en experimentos de complementación se agruparon en un único clado (Fig. 4 y Tabla 3). La identidad de secuencia de las proteínas de este clado oscila entre del 33% al 100% (Tabla 3). El clado DRL está respaldado por valores de "bootstrap" relativamente bajos (64), pero se obtuvo consistentemente empleando otros dos métodos independientes de reconstrucción filogenética ("Neighbour Joining" y "Maximum Parsimony"). Todos los organismos de clase A tienen al menos una secuencia de DRL en este clado (Tabla 4). La única excepción es la proteína B5J045 de la cepa 307 de la alfa-proteobacteria *Octadecabacter antarcticus*. aunque en el genoma de la cepa 238 existe una secuencia DRL incluida en el clado (B5K941) (Tabla 4). El único organismo de clase C con una secuencia DRL en el clado es *Chloroflexus aurantiacus* (Tabla 4). El resto de secuencias del clado son de organismos de clase B (i.e., que presentan una secuencia DXR en sus genomas, aunque no necesariamente activa enzimáticamente; Tabla 2). La mayoría de estos

organismos eran firmicutes pero también están incluidas en el clado secuencias de la alfa-proteobacteria *Roseobacter litoralis*, la beta-proteobacteria *Verminephrobacter eiseniae* y la actinobacteria marina PHSC20C1 (Fig. 4 y Tabla 4). Todas las secuencias de organismos de clase B excluidas del clado DRL que se ensayaron en experimentos de complementación resultaron ser inactivas (Tabla 2), lo que indica que no eran enzimas DRL auténticas.

Generación de un perfil HMM a partir del análisis de las diez secuencias

Además, por otra parte, se realizó un alineamiento de las diez secuencias aminoacídicas verificadas como DRL activas (Tabla 2) para generar un perfil de HMM usando el programa HMMER (<http://HMMER.janelia.org/>). El perfil de HMM fue usado para modelar las DRL y en búsquedas de posibles homólogos de DRL en grandes bases de datos de proteínas como Pfam o UNIPROT (cf. R. Apweiler *et al.*, “UniProt: the Universal Protein knowledgebase”, *Nucleic Acids Res* 2004. vol. 32, pp. 115-119). El protocolo completo aparece detallado a continuación:

(i) Alineamiento múltiple de secuencias de proteínas DRL. Las secuencias fueron alineadas con el programa CLUSTALW usando los parámetros por defecto (cf. J.D. Thompson *et al.*, “CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice” *Nucleic Acids Res* 1994, vol. 22, pp. 4673-4680). Los porcentajes de identidad a nivel de secuencia con respecto a la proteína DRL de *B. abortus* (Q2YIM3) eran altamente variables, oscilando entre un 90% (A6WYQ0; *O. anthropi*) y un 33% (B0S038; *F. magna*). Estos datos proporcionan una estimación de la distancia genética y evolutiva entre las secuencias alineadas, revelando la elevada divergencia evolutiva entre las DRL funcionales.

(ii) Construcción de un perfil de HMM usando el programa hmmbuild del paquete de aplicaciones HMMER, v. 2.3.3. El programa hmmbuild lee un fichero en formato FASTA con el alineamiento de secuencias DRL, construye un perfil de HMM y lo graba en un fichero de salida.

(iii) Calibración del perfil de HMM usando “hmmcalibrate” del paquete de aplicaciones HMMER, v. 2.3.3.

(iv) Examen de la especificidad y sensibilidad del perfil de HMM en búsquedas de secuencias DRL: El perfil de HMM, una vez calibrado, fue usado como “query” para buscar secuencias DRL. Se usó para ello el programa “hmmsearch” del paquete de aplicaciones HMMER, v. 2.3.3. El programa “hmmsearch” lee un fichero HMMER con un perfil de HMM, busca en una base de datos secuencias de proteínas con similitudes significativas y devuelve un listado de las secuencias resultantes ordenadas. La base de datos utilizada fue UNIPROT. Entre los mejores resultados aparecían las diez secuencias de DRL funcionales usadas para generar el perfil, con puntuaciones/valores-E de 860/0 (Q2YIM3; *B. abortus*) a 267/2.00E-69 (B0S038; *F. magna*). Sin embargo, seis secuencias adicionales que también mostraban puntuaciones/valores-E significativos, desde 50/3E-04 (B4EB12, *B. cepacia*) a 243/2E-62 (B2IVC2, *N. punctiforme*). no consiguieron complementar el mutante de *E. coli dxr* defectivo EcAB4-10, lo que indicaba que no poseían actividad enzimática DRL. El programa “hmmsearch” fue también usado para buscar entre las secuencias de proteína incluidas en las bases de datos Pfam no encontrándose ninguna con puntuaciones/valores-E significativos.

Estos resultados indican que (i) las DRLs funcionales comparten similitud a nivel de secuencia significativa; (ii) los perfiles de HMM muestran gran sensibilidad para identificar con seguridad a los homólogos de DRL funcionales en las búsquedas en INTERPRO; y (iii) el perfil de HMM muestra suficiente especificidad para descartar aquellas secuencias previamente clasificadas en las familias de proteínas de Pfam. En conclusión, aunque el perfil de HMM funciona bien para identificar hipotéticos homólogos de DRL, algunas de las secuencias identificadas parecen corresponder a homólogos divergidos funcionalmente para los que la actividad enzimática DRL se ha perdido o bien no ha sido adquirida a través de la evolución.

Para refinar la definición de DRL, se sometió a todas las secuencias con homología con DRL (Tabla 3) a un análisis filogenético a partir de alineamientos de secuencia de proteína obtenidos mediante CLUSTALW. Se usaron tres métodos de reconstrucción filogenética: máxima verosimilitud, máxima parsimonia y “neighbor-joining”. Los análisis de verosimilitud fueron llevados a cabo con el programa PHYML v2.4.5, usando el modelo JTT de evolución de proteínas (cf. D.T. Jones *et al.*, *supra*; S. Guindon and O. Gascuel, *supra*). Los análisis de “neighbor-joining” y máxima parsimonia se realizaron en MEGA 4.0, usando la configuración por defecto en cada caso (cf. K. Tamura *et al.*, *supra*). Para dotar de confianza estadística a la topología del árbol obtenido, se llevó a cabo un análisis de “bootstrap” con 1000 replicaciones (cf. J. Felsenstein, “Confidence Limits on Phylogenies: An Approach Using the Bootstrap” *Evolution* 1985. vol. 39, pp. 783-791).

Cuando las secuencias aminoacídicas identificadas a partir del perfil de HMM generado o a partir del BLASTp fueron sujetas a análisis filogenéticos por el método de máxima verosimilitud, se observó que todas las secuencias DRL que habían mostrado ser funcionales en los ensayos de complementación agrupaban en un grupo monofilético o clado (Fig. 4). Dicho clado mostraba valores de “bootstrap” relativamente bajos, pero era sistemáticamente obtenido por los dos métodos alternativos de reconstrucción filogenética que fueron empleados. De acuerdo a la teoría filogenética, las secuencias que agrupan en un clado son homologas y, por lo tanto, compartirían un ancestro común. Coherentemente, todas las secuencias fuera del clado DRL que fueron examinadas experimentalmente no fueron capaces de complementar el mutante EcAB4-10 (Tabla 2). Estos resultados indican que (i) las secuencias que agrupan dentro del clado DRL pertenecen a la familia DRL y comparten actividad enzimática DRL y (ii) las secuencias que muestran identidad de secuencia significativa con las DRL pero agrupan fuera del clado DRL no serían enzimas DRL funcionales.

ES 2 372 942 A1

TABLA 3

La columna “clase” indica la clasificación de las cepas bacterianas según la distribución de secuencias con homología a DRL, DXR y/o otras enzimas de la vía MEP en la misma cepa: A (+DRL, -DXR, +MEP), B (+DRL, +DXR, +MEP), C (+DRL, -DXR, -MEP). ^(a)únicamente se indican los hits anotados como DXS. ^(b)indica enzima bifuncional MCT/MDS (IspD/IspF). La columna “DRL clado” indica las secuencias que pertenecen (+) o no (-) al clado DRL mostrado en la Fig. 4 Las secuencias DRL múltiples en la misma cepa están separadas por barras. La columna “C-DRL” indica las secuencias que complementan (+) o no (-) el mutante dxr *E. coli*. Las cepas de las que se ha analizado experimentalmente su secuencia están marcadas en negrita.

Clase	Organismo	Número de secuencias							DRL clado	C-DRL	
		via MEP									
		DXS ^a	DXR	MCT	CMK	MDS	HDS	HDR			
	Bartonella bacilliformis (strain ATCC 35685 / KC583)	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Bartonella henselae	1	0	1 ^b	1	1 ^b	1	1	1	+	+
	Bartonella quintana	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Bartonella tribocorum (strain CIP 105476 / IBS 506)	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Brucella abortus	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Brucella abortus 2308	1	0	1 ^b	1	1 ^b	1	1	1	+	+
	Brucella abortus S19	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Brucella canis ATCC 23365	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Brucella melitensis 16M	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Brucella ovis 63/290	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Brucella suis 1330	1	0	1 ^b	1	1 ^b	1	1	1	+	
A	Brucella suis ATCC 23445	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Fulvimarina pelagi HTCC2506	1	0	1 ^b	1	1 ^b	1	1	2	+/-	
	Aurantimonas sp. SI85-9A1	1	0	1 ^b	1	1 ^b	1	1	3	+/-/-	
	Mesorhizobium sp. BNC1	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Ochrobactrum anthropi LMG3331	1	0	1 ^b	1	1 ^b	1	1	2	+/+	+/+
	Octadecabacter antarcticus 238	1	0	1 ^b	1	1 ^b	1	1	1	+	
	Octadecabacter antarcticus 307	1	0	1 ^b	1	1 ^b	1	1	1	-	
	Mesorhizobium loti MAFF303099	1	0	1 ^b	1	1 ^b	1	1	2	+/+	+/+
	Anaerococcus hydrogenalis DSM 7454	1	0	1	1	1	1	1	1	+	
	Anaerofustis stercorihominis DSM 17244	1	0	2	1	1	1	1	2	+/+	
	Fingoldia magna ATCC 29328	0	0	1	1	1	1	1	1	+	+
	Parvimonas micra ATCC 33270	1	0	2	1	1	1	1	1	+	
B	Arthrobacter sp. (strain FB24)	1	1	1	1	1	1	1	1	-	
	Kineococcus radiotolerans (strain ATCC BAA-149 / DSM 14245 / SRS30216)	1	2	1	1	1	1	1	1	-	
	marine actinobacterium PHSC20C1	1	1	1 ^b	1	1 ^b	1	1	1	+	
	Mycobacterium smegmatis mc(2)155	1	1	1	1	1	1	1	1	-	-
	Streptomyces griseus subsp. griseus (strain JCM 4626 / NBRC 13350)	1	1	2	1	1	1	1	2	+/-	
	Acidiphilium cryptum (strain JF-5)	1	1	1	1	1	1	1	1	-	
	Agrobacterium tumefaciens C58	1	1	1 ^b	1	1 ^b	1	1	1	-	-
	alpha proteobacterium BAL199	1	1	1 ^b	1	1 ^b	1	1	1	-	
	Azorhizobium caulinodans (strain ATCC 43989 / DSM 5975 / ORS 571)	1	1	1 ^b	1	1 ^b	1	1	1	-	
	Bradyrhizobium japonicum	1	1	1 ^b	1	1 ^b	1	1	1	-	
	Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	1	1	1 ^b	1	1 ^b	1	1	1	-	
	Bradyrhizobium sp. (strain ORS278)	1	1	1 ^b	1	1 ^b	1	1	1	-	
	Candidatus Pelagibacter sp. HTCC7211	0	1	1 ^b	1	1 ^b	1	0	1	-	
	Candidatus Pelagibacter ubique	1	1	1 ^b	1	1 ^b	1	1	1	-	-

ES 2 372 942 A1

HTCC1002										
	Hoeflea phototrophica DFL-43	1	1	1 ^b	1	1 ^b	1	1	2	-/-
5	Magnetospirillum gryphiswaldense	0	1	1 ^b	1	1 ^b	1	0	1	-
	Methylobacterium nodulans ORS 2060	2	1	1 ^b	1	1 ^b	1	1	2	-/-
	Methylobacterium radiotolerans (strain ATCC 27329 / DSM 1819 / JCM 2831)	2	1	1 ^b	1	1 ^b	1	1	1	-
10	Methylobacterium sp. (strain 4-46)	2	1	1 ^b	1	1 ^b	1	1	2	-/-
	Paracoccus denitrificans (strain Pd 1222)	1	1	1 ^b	1	1 ^b	1	1	1	-
15	Pelagibacter ubique	1	1	1 ^b	1	1 ^b	1	1	1	-
	Rhodobacter sphaeroides (strain ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158)	2	1	1	1	1	1	1	1	-
20	Rhodobacter sphaeroides (strain ATCC 17025 / ATH 2.4.3)	2	1	1	1	1	1	1	1	-
	Rhodobacter sphaeroides (strain ATCC 17029 / ATH 2.4.9)	2	1	1	1	1	1	1	1	-
25	Rhodopseudomonas palustris (strain BisB5)	1	1	1 ^b	1	1 ^b	1	1	1	-
	Rhodospirillum rubrum (strain ATCC 11170 / NCIB 8255)	2	1	1 ^b	1	1 ^b	1	1	1	-
	Roseobacter litoralis Och 149	2	1	1	1	1	1	1	1	+
30	Roseobacter sp. AzwK-3b	2	1	1 ^b	1	1 ^b	1	1	1	-
	Roseovarius sp. 217	2	1	1 ^b	1	1 ^b	1	1	1	-
	Sphingomonas wittichii (strain RW1 / DSM 6014 / JCM 10273)	1	1	1 ^b	1	1 ^b	1	1	1	-
35	Acidovorax avenae subsp. citrulli (strain AAC00-1)	1	1	1	1	1	1	1	1	-
	Acidovorax sp. (strain JS42)	1	1	1	1	1	1	1	2	-/-
	Azoarcus sp. (strain BH72)	1	1	1	1	1	1	1	1	-
40	Azoarcus sp. (strain EbN1) (Aromatoleum aromaticum (strain EbN1))	1	1	1	1	1	1	1	1	-
	Bordetella avium (strain 197N)	1	1	1	1	1	1	1	1	-/-
	Bordetella bronchiseptica (Alcaligenes bronchisepticus)	1	1	1	1	1	1	1	1	-
45	Bordetella parapertussis	1	1	1	1	1	1	1	1	-
	Bordetella petrii (strain ATCC BAA-461 / DSM 12804 / CCUG 43448)	1	1	1	1	1	1	1	2	-/-
50	Burkholderia ambifaria (strain ATCC BAA-244 / AMMD) (B. cepacia (strain AMMD))	1	1	1	1	1	1	1	1	-
	Burkholderia ambifaria (strain MC40-6)	1	1	1	1	1	1	1	1	-
55	Burkholderia ambifaria IOP40-10	1	1	1	1	1	1	1	1	-
	Burkholderia ambifaria MEX-5	1	1	0	1	1	1	1	1	-
	Burkholderia cenocepacia (strain AU 1054)	1	1	1	1	1	1	1	1	-
60	Burkholderia cenocepacia (strain HI2424)	1	1	1	1	1	1	1	1	-
	Burkholderia cenocepacia (strain MC0-3)	1	1	1	1	1	1	1	1	-

65

ES 2 372 942 A1

	Burkholderia cenocepacia PC184	1	1	1	1	1	1	1	1	-
	Burkholderia cepacia J2315	1	1	1	0	1	1	1	1	-
5	Burkholderia dolosa AUO158	1	1	0	1	1	1	1	1	-
	Burkholderia graminis C4D1M	1	1	1	1	1	1	1	1	-
	Burkholderia multivorans (strain ATCC 17616 / 249)	1	1	1	1	1	1	1	1	-
10	Burkholderia phymatum (strain DSM 17167 / STM815)	1	1	1	1	1	1	1	1	-
	Burkholderia phytofirmans (strain DSM 17436 / PsJN)	1	1	1	1	1	1	1	1	-
15	Burkholderia sp. (strain 383) (B. cepacia (strain ATCC 17760 / NCIB 9086 / R18194))	1	1	1	1	1	1	1	1	-
	Burkholderia sp. H160	1	1	1	1	1	1	1	1	-
20	Burkholderia vietnamiensis (strain G4 / LMG 22486) (B. cepacia (strain R1808))	1	1	1	1	1	1	1	1	-
	Burkholderia xenovorans (strain LB400)	1	1	1	1	1	1	1	1	-
	Chromobacterium violaceum	1	1	1	1	1	1	1	1	-
25	Cupriavidus taiwanensis (strain R1 / LMG 19424) (Ralstonia taiwanensis)	1	1	1	1	1	1	1	2	-/-
	Dechloromonas aromatica (strain RCB)	1	1	1	1	1	1	1	1	-
30	Janthinobacterium sp. (strain Marseille) (Minibacterium massiliensis)	1	1	1	1	1	1	1	1	-
	Leptothrix cholodnii ATCC 51168 / LMG 8142 / SP-6 (L. discophora SP-6)	1	1	1	1	1	1	1	1	-
35	Limnobacter sp. MED105	1	1	1	1	1	1	1	1	-
	Methylibium petroleiphilum (strain PM1)	1	1	1	1	1	1	1	2	-/-
40	Methylobacillus flagellatum (Methylobacillus flagellatus)	1	1	1	1	1	1	1	1	-
	Methylobacillus flagellatus (strain KT / ATCC 51484 / DSM 6875)	1	1	1	1	1	1	1	1	-
45	Neisseria gonorrhoeae (strain NCCP11945)	1	1	1	1	1	1	1	1	-
	Nitrosomonas eutropha (strain C91)	1	1	1	1	1	1	1	1	-
	Nitrospira multiformis (strain ATCC 25196 / NCIMB 11849)	1	1	1	1	1	1	1	1	-
50	Polaromonas naphthalenivorans (strain CJ2)	1	1	1	1	1	1	1	1	-
	Polaromonas sp. (strain JS666 / ATCC BAA-500)	1	1	1	1	1	1	1	2	-/-
55	Polynucleobacter necessarius (strain STIR1)	1	1	1	1	1	1	1	1	-
	Polynucleobacter sp. (strain QLW-P1DMWA-1)	1	1	1	1	1	1	1	1	-
60	Ralstonia eutropha ATCC 17699 / H16 / DSM 428 / Stanier 337 (Cupriavidus necator)	1	1	1	1	1	1	1	1	-
	Ralstonia eutropha (strain JMP134) (Alcaligenes eutrophus)	1	1	1	1	1	1	1	1	-

ES 2 372 942 A1

	Ralstonia metallidurans (strain CH34 / ATCC 43123 / DSM 2839)	1	1	1	1	1	1	1	1	-	
5	Ralstonia pickettii (strain 12J)	1	1	1	1	1	1	1	1	-	
	Ralstonia pickettii 12D	1	1	1	1	1	1	1	1	-	
	Ralstonia solanacearum (Pseudomonas solanacearum)	1	1	1	1	1	1	1	2	-/-	
10	Ralstonia solanacearum IPO1609	1	1	1	1	1	1	1	1	-	
	Thiobacillus denitrificans (strain ATCC 25259)	1	1	1	1	1	1	1	1	-	
	Variovorax paradoxus S110	1	1	1	1	1	1	1	3	-/-/-	
15	Verminephrobacter eiseniae (strain EF01-2)	1	1	1	1	1	1	1	1	+	
	Anabaena sp. (strain PCC 7120)	1	1	1	1	1	1	1	1	-	
	Anabaena variabilis PCC 7937	1	1	1	1	1	1	1	1	-	-
	Cyanothece sp. PCC 7425	1	1	1	1	1	1	1	1	-	
20	Gloeobacter violaceus	1	1	1	1	1	1	1	2	-/-	
	Nodularia spumigena CCY 9414	1	1	1	1	1	1	1	2	-/-	
	Nostoc punctiforme PCC 73102	1	1	1	1	1	1	1	1	-	-
	Synechococcus sp. PCC 7335	1	1	1	1	1	1	1	1	-	
25	Anaeromyxobacter sp. (strain Fw109-5)	1	1	1 ^b	1	1 ^b	1	1	1	-	
	Desulfococcus oleovorans (strain DSM 6200 / Hxd3)	1	1	1	1	1	1	1	1	-	
30	Desulfuromonas acetoxidans DSM 684	1	1	1	1	1	1	1	1	-	
	Anaerostipes caccae DSM 14662	1	1	1	1	1	1	1	1	+	
	Bacillus clausii (strain KSM-K16)	1	1	1	1	1	1	1	1	+	
35	Bacillus halodurans C-125	1	1	1	1	1	1	1	1	+	+
	Bacillus pumilus (strain SAFR-032)	1	1	1	1	1	1	1	1	+	
	Bacillus pumilus ATCC 7061	1	1	1	1	1	1	1	1	+	
	Clostridium bartlettii DSM 16795	1	1	1	1	1	1	1	1	+	
40	Clostridium bolteae ATCC BAA-613	1	1	2	1	1	1	1	1	-	
	Clostridium difficile (strain 630)	1	1	1	1	1	1	1	1	+	
	Listeria innocua	1	1	2	1	1	0	0	1	+	
	Listeria monocytogenes	1	1	1	1	1	1	1	1	+	
45	Listeria monocytogenes HCC23	1	1	2	1	1	0	0	1	+	
	Listeria monocytogenes serotype 4b (strain F2365)	1	1	2	1	1	1	1	1	+	+
	Listeria monocytogenes str. 1/2a F6854	1	1	2	1	1	1	1	1	+	
50	Listeria monocytogenes P14	1	1	1	1	0	1	1	1	+	
	Listeria welshimeri serovar 6b (strain ATCC 35897 / DSM 20650 / SLCC5334)	1	1	2	1	0	0	0	1	+	
55	Paenibacillus sp. JDR-2	1	1	1	1	1	1	1	1	-	
	Beggiatoa sp. PS	1	1	1	1	1	1	1	1	-	
	Marinobacter algicola DG893	1	1	1	1	1	1	1	1	-	
	Marinobacter sp. ELB17	1	1	1	1	1	1	1	1	-	
60	Marinomonas sp. (strain MWYL1)	1	1	1	1	1	1	1	1	-	
	Methylococcus capsulatus	1	1	1	1	1	1	1	1	-	
	Moritella sp. PE36	1	1	1	1	1	1	1	1	-	
	Oceanobacter sp. RED65	1	1	1	1	1	1	1	1	-	
65	Pseudomonas aeruginosa	1	1	1	1	1	1	1	1	-	

ES 2 372 942 A1

	Pseudomonas aeruginosa (strain UCBPP-PA14)	1	1	1	1	1	1	1	1	-
5	Pseudomonas aeruginosa 2192	1	1	0	1	1	1	1	1	-
	Pseudomonas aeruginosa C3719	1	1	0	1	1	1	1	1	-
	Pseudomonas aeruginosa LESB58	1	1	1	1	1	1	1	1	-
	Pseudomonas fluorescens (strain Pf-5 / ATCC BAA-477)	1	1	1	1	1	1	1	1	-
10	Pseudomonas fluorescens (strain Pf0-1)	1	1	1	1	1	1	1	1	-
	Pseudomonas stutzeri (strain A1501)	1	1	1	1	1	1	1	1	-
	Xanthobacter autotrophicus (strain ATCC BAA-1158 / Py2)	1	1	1 ^b	1	1 ^b	1	1	1	-
15	Xanthomonas campestris pv. campestris	1	1	1	1	1	1	1	1	-
	Xanthomonas campestris pv. campestris (strain 8004)	1	1	1	1	1	1	1	1	-
20	Xanthomonas campestris pv. campestris (strain B100)	1	1	1	1	1	1	1	1	-
	Candidatus Methanosphaerula palustris E1-9c	0	0	0	0	0	0	0	1	-
25	uncultured marine microorganism HF4000_005K23	0	0	0	0	0	0	0	1	-
	Streptomyces glaucescens	0	0	0	0	0	0	0	1	-
	Paracoccus methylutens	0	0	0	0	0	0	0	1	-
	Methylobacillus glycogenes	0	0	0	0	0	0	0	1	-
30	Coxiella burnetii RSA 334	0	0	0	0	0	0	0	1	-
	Coxiella burnetii	0	0	0	0	0	0	0	1	-
	C Coxiella burnetii (strain CbuG_Q212) (Coxiella burnetii (strain Q212))	0	0	0	0	0	0	0	1	-
35	Coxiella burnetii (strain CbuK_Q154) (Coxiella burnetii (strain Q154))	0	0	0	0	0	0	0	1	-
	Coxiella burnetii (strain Dugway 5J108-111)	0	0	0	0	0	0	0	1	-
40	Coxiella burnetii (strain RSA 331 / Henzerling II)	0	0	0	0	0	0	0	1	-
	Chloroflexus aggregans DSM 9485	0	0	0	0	0	0	0	2	-/-
45	Chloroflexus aurantiacus (strain ATCC 29366 / DSM 635 / J-10-fl)	0	0	0	0	0	0	0	3	+/-/-

50

55

60

65

ES 2 372 942 A1

TABLA 4

Los "E-value" son relativos a la proteína DRL de *B. abortus* (Q2YIM3). La columna "DRL clado" indica las secuencias que pertenecen (+) o no (-) al clado DRL mostrado en la Fig. 4. La columna "C-DRL" indica las secuencias que complementan (+) o no (-) el mutante deficiente en DXR de *E. coli* (ver también Tabla 2). Las secuencias probadas y las correspondientes cepas se marcan en negrita. La columna "clase" indica la clasificación de las cepas bacterianas según la distribución de las secuencias con homología a DRL, DXR y/o otras enzimas de la vía MEP en la misma cepa: A (+DRL, -DXR, +MEP), B (+DRL, +DXR, +MEP), C (+DRL, -DXR, -MEP).

Organismo	Accession	E-value	DRL clado	C-DRL	Clase
Brucella abortus (strain 2308)	Q2YIM3	0	+	+	A
Brucella abortus	Q579I6	0	+		A
Brucella abortus (strain S19)	B2SD59	0	+		A
Brucella canis (strain ATCC 23365 / NCTC 10854)	A9MCQ0	0	+		A
Brucella melitensis	Q8YD51	0	+		A
Brucella ovis (strain ATCC 25840 / 63/290 / NCTC 10512)	A5VVN1	0	+		A
Brucella suis	Q8FV77	0	+		A
Brucella suis (strain ATCC 23445 / NCTC 10510)	A9WVY0	0	+		A
Mesorhizobium sp. (strain BNC1)	Q11C19	0	+		A
Ochrobactrum anthropi (ATCC 49188 / DSM 6882 / NCTC 12168)	A6WYQ0	0	+	+	A
Ochrobactrum anthropi (ATCC 49188 / DSM 6882 / NCTC 12168)	A6X6G6	0	+	+	A
Rhizobium loti (Mesorhizobium loti)	Q98FT2	0	+	+	A
Bartonella henselae (Rochalimaea henselae)	Q6G2D9	1,00E-162	+	+	A
Bartonella tribocorum (strain CIP 105476 / IBS 506)	A9IWX4	1,00E-162	+		A
Bartonella bacilliformis (strain ATCC 35685 / KC583)	A1UTR5	1,00E-161	+		A
Bartonella quintana (Rochalimaea quintana)	Q6FYY9	1,00E-161	+		A
Rhizobium loti (Mesorhizobium loti)	Q989B6	1,00E-161	+	+	A
Verminephrobacter eiseniae (strain EF01-2)	A1WK22	1,00E-157	+		B
Octadecabacter antarcticus 238	B5K941	1,00E-151	+		A
Roseobacter litoralis Och 149	A9HDV1	1,00E-151	+	+	B
Fulvimarina pelagi HTCC2506	Q0G7L0	1,00E-123	+		A
Manganese-oxidizing bacterium (strain SI85-9A1)	Q1YFT8	1,00E-123	+		A
Bacillus halodurans	Q9KES5	6,00E-93	+	+	B
Anaerostipes caccae DSM 14662	B0MHR8	4,00E-88	+		B
Clostridium difficile (strain 630)	Q182A2	4,00E-88	+		B
Bacillus clausii (strain KSM-K16)	Q5WCY8	9,00E-88	+		B
marine actinobacterium PHSC20C1	A4AIF9	5,00E-87	+		B
Bacillus pumilus (strain SAFR-032)	A8FCZ5	6,00E-84	+		B
Bacillus pumilus ATCC 7061	B4AE91	1,00E-83	+		B
Anaerofustis stercorihominis DSM 17244	B1C6E7	1,00E-81	+		A
Chloroflexus aurantiacus (ATCC 29366 / DSM 635 / J-10-fl)	A9W9R9	4,00E-81	+		C
Anaerofustis stercorihominis DSM 17244	B1CB96	5,00E-81	+		A
Listeria welshimeri serovar 6b (ATCC 35897 / DSM 20650 / SLCC5334)	A0AG02	3,00E-80	+		B
Listeria innocua	Q92EB0	2,00E-79	+		B
Clostridium bartlettii DSM 16795	B0A6H1	3,00E-79	+		B

ES 2 372 942 A1

	Listeria monocytogenes HCC23	B8DA79	5,00E-79	+		B
	Listeria monocytogenes	Q8Y9I2	8,00E-79	+		B
5	Listeria monocytogenes serotype 4b (strain F2365)	Q723A4	8,00E-79	+	+	B
	Listeria monocytogenes P14	Q4EIE8	8,00E-79	+		B
	Listeria monocytogenes str. 1/2a F6854	Q4EVU9	1,00E-78	+		B
10	Finegoldia magna (ATCC 29328) (Peptostreptococcus magnus)	B0S038	2,00E-69	+	+	A
	Parvimonas micra ATCC 33270	A8SLM3	3,00E-69	+		A
15	Rhodobacter sphaeroides (strain ATCC 17025 / ATH 2.4.3)	A4WUK5	6,00E-66	-		B
	Chloroflexus aggregans DSM 9485	B8GCI5	3,00E-65	-		C
	Chloroflexus aurantiacus (strain ATCC 29366 / DSM 635 / J-10-fl)	A9WJZ1	9,00E-64	-		C
20	Gloeobacter violaceus	Q7NIH9	4,00E-63	-		B
	Rhodobacter sphaeroides (ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158)	Q3J1Z1	4,00E-63	-		B
25	Rhodobacter sphaeroides (strain ATCC 17029 / ATH 2.4.9)	A3PK98	5,00E-63	-		B
	Nostoc punctiforme (strain ATCC 29133 / PCC 73102)	B2IVC2	2,00E-62	-	-	B
30	Xanthomonas campestris pv. campestris (strain B100)	B0RT81	2,00E-62	-		B
	Gloeobacter violaceus	Q7NII0	7,00E-62	-		B
	Anaeromyxobacter sp. (strain Fw109-5)	A7HEZ6	3,00E-61	-		B
35	Limnobacter sp. MED105	A6GTN0	3,00E-61	-		B
	Anaerococcus hydrogenalis DSM 7454	B6W7N4	4,00E-61	+		A
	Cyanothece sp. PCC 7425	B8HQ87	2,00E-60	-		B
	Nodularia spumigena CCY 9414	A0ZLW5	1,00E-59	-		B
40	Anabaena sp. (strain PCC 7120)	Q8YT93	4,00E-59	-		B
	Acidiphilium cryptum (strain JF-5)	A5FYT3	2,00E-58	-		B
	Anabaena variabilis (strain ATCC 29413 / PCC 7937)	Q3ME48	5,00E-58	-	-	B
45	Acidovorax sp. (strain JS42)	A1W8A1	2,00E-57	-		B
	Agrobacterium tumefaciens (strain C58 / ATCC 33970)	A9CES2	2,00E-57	-	-	B
	Chloroflexus aggregans DSM 9485	B8G2T8	1,00E-56	-		C
50	Rhodopseudomonas palustris (strain BisB5)	Q131D1	2,00E-56	-		B
	Methylobacterium radiotolerans (ATCC 27329 / DSM 1819 / JCM 2831)	B1LSV9	6,00E-56	-		B
55	Sphingomonas wittichii (strain RW1 / DSM 6014 / JCM 10273)	A5VF13	6,00E-56	-		B
	Streptomyces griseus subsp. griseus (strain JCM 4626 / NBRC 13350)	Q53IE8	6,00E-56	-		B
60	Streptomyces griseus subsp. griseus (strain JCM 4626 / NBRC 13350)	B1W361	6,00E-56	-		B
	Bradyrhizobium japonicum	Q89CI2	7,00E-56	-		B
65	Kineococcus radiotolerans (ATCC BAA-149 / DSM 14245 / SRS30216)	A6W5H2	9,00E-56	-		B
	Fulvimarina pelagi HTCC2506	Q0G3F9	1,00E-55	-		A

ES 2 372 942 A1

	Bradyrhizobium sp. (strain ORS278)	A4Z1N0	4,00E-55	-	B
	Variovorax paradoxus S110	B7YYT2	5,00E-55	-	B
5	Bordetella bronchiseptica (Alcaligenes bronchisepticus)	Q7WQA1	6,00E-55	-	B
	Xanthomonas campestris pv. campestris	Q8P954	6,00E-55	-	B
	Xanthomonas campestris pv. campestris (strain 8004)	Q4UUP5	6,00E-55	-	B
10	Paracoccus methylutens	A7UBP4	1,00E-54	-	C
	Polaromonas sp. (strain JS666 / ATCC BAA-500)	Q12CU3	1,00E-54	-	B
15	Bordetella parapertussis	Q7WC97	2,00E-54	-	B
	Chloroflexus aurantiacus (strain ATCC 29366 / DSM 635 / J-10-fl)	A9WI90	2,00E-54	-	C
	Bradyrhizobium sp. (strain BTAi1 / ATCC BAA-1182)	A5EB98	5,00E-54	-	B
20	Hoeflea phototrophica DFL-43	A9DBL9	2,00E-53	-	B
	Manganese-oxidizing bacterium (strain SI85-9A1)	Q1YMW1	2,00E-53	-	A
25	Streptomyces glaucescens	Q54201	8,00E-53	-	C
	Bordetella petrii (strain ATCC BAA-461 / DSM 12804 / CCUG 43448)	A9IDL3	2,00E-52	-	B
	Methylobacterium nodulans ORS 2060	B8IR91	4,00E-52	-	B
30	Azorhizobium caulinodans (strain ATCC 43989 / DSM 5975 / ORS 571)	A8I2L2	5,00E-52	-	B
	alpha proteobacterium BAL199	A8TV43	6,00E-52	-	B
	Polynucleobacter sp. (strain QLW-P1DMWA-1)	A4SZ50	2,00E-51	-	B
35	Methylbium petroleiphilum (strain PM1)	A2SKW3	3,00E-51	-	B
	Methylobacterium sp. (strain 4-46)	B0UJP5	3,00E-51	-	B
	Methylobacterium nodulans ORS 2060	B8IHI2	9,00E-51	-	B
40	Methylobacterium sp. (strain 4-46)	B0UNE5	2,00E-50	-	B
	Roseobacter sp. AzwK-3b	A6FTB1	3,00E-50	-	B
	Roseovarius sp. 217	A3W1Q2	3,00E-50	-	B
	Octadecabacter antarcticus 307	B5J045	7,00E-49	-	A
45	Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155)	A0QQV9	4,00E-48	-	B
	Xanthobacter autotrophicus (strain ATCC BAA-1158 / Py2)	A7IDN8	4,00E-48	-	B
50	Hoeflea phototrophica DFL-43	A9DBD0	5,00E-47	-	B
	Arthrobacter sp. (strain FB24)	A0K110	1,00E-46	-	B
	Cupriavidus taiwanensis (R1 / LMG 19424) (Ralstonia taiwanensis)	B3RD23	1,00E-46	-	B
55	Desulfococcus oleovorans (strain DSM 6200 / Hxd3)	A8ZXQ6	3,00E-46	-	B
	Bordetella avium (strain 197N)	Q2KZS1	2,00E-44	-	B
	Coxiella burnetii (strain RSA 331 / Henzerling II)	A9NAT2	1,00E-43	-	C
60	Moritella sp. PE36	A6FB13	1,00E-43	-	B
	Coxiella burnetii	Q83AP5	2,00E-43	-	C
	Coxiella burnetii (strain CbuK_Q154) (Coxiella burnetii (strain Q154))	B6J9D4	2,00E-43	-	C
65	Coxiella burnetii (strain Dugway 5J108-	A9KF08	2,00E-43	-	C

ES 2 372 942 A1

	111)				
5	Coxiella burnetii (strain CbuG_Q212) (Coxiella burnetii (strain Q212))	B6J325	4,00E-43	-	C
	Coxiella burnetii RSA 334	A9ZG83	5,00E-43	-	C
	Manganese-oxidizing bacterium (strain SI85-9A1)	Q1YHS2	6,00E-42	-	A
10	Candidatus Pelagibacter ubique HTCC1002	Q1V2P9	1,00E-39	-	B
	Pelagibacter ubique	Q4FMC1	3,00E-39	-	B
	Magnetospirillum gryphiswaldense	A4U0P3	1,00E-38	-	B
15	Candidatus Pelagibacter sp. HTCC7211 uncultured marine microorganism HF4000_005K23	B6BT86	3,00E-38	-	B
		B3T0M6	4,00E-36	-	C
20	Rhodospirillum rubrum (strain ATCC 11170 / NCIB 8255)	Q2RPK9	1,00E-28	-	B
	Pseudomonas stutzeri (strain A1501)	A4VLJ1	6,00E-28	-	B
	Paracoccus denitrificans (strain Pd 1222)	A1B7G7	5,00E-27	-	B
	Clostridium bolteae ATCC BAA-613	A8RH36	7,00E-27	-	B
25	Candidatus Methanosphaerula palustris E1-9c	B8GEI6	2,00E-26	-	C
	Methylobacillus flagellatus (strain KT / ATCC 51484 / DSM 6875)	Q1H2H1	7,00E-08	-	B
30	Methylobacillus flagellatum	Q9RAM9	1,00E-07	-	B
	Methylobacillus glycozenes	P37144	1,00E-07	-	C
	Chromobacterium violaceum	Q7NZC6	4,00E-07	-	B
35	Burkholderia phymatum (strain DSM 17167 / STM815)	B2JG38	3,00E-06	-	B
	Burkholderia xenovorans (strain LB400)	Q13ZA0	3,00E-06	-	B
	Nitrosospira multiformis (strain ATCC 25196 / NCIMB 11849)	Q2Y8R7	3,00E-06	-	B
40	Desulfuromonas acetoxidans DSM 684	Q1K0M5	7,00E-06	-	B
	Burkholderia sp. H160	B5WNF8	9,00E-06	-	B
	Azoarcus sp. (strain BH72)	A1K793	3,00E-05	-	B
	Burkholderia graminis C4D1M	B1FYR6	3,00E-05	-	B
45	Burkholderia phytofirmans (strain DSM 17436 / PsJN)	B2T3P9	3,00E-05	-	B
	Burkholderia sp. 383 (B. cepacia ATCC 17760 / NCIB 9086 / R18194))	Q39FL7	3,00E-05	-	B
50	Burkholderia vietnamiensis G4 / LMG 22486 (B. cepacia R1808)	A4JET2	3,00E-05	-	B
	Ralstonia eutropha ATCC 17699 / H16 / DSM 428 / Stanier 337 (Cupriavidus necator)	Q0K9F8	3,00E-05	-	B
55	Ralstonia eutropha (strain JMP134) (Alcaligenes eutrophus)	Q46ZS6	3,00E-05	-	B
	Methylibium petroleiphilum (strain PM1)	A2SG14	4,00E-05	-	B
60	Beggiatoa sp. PS	A7BUQ0	6,00E-05	-	B
	Burkholderia cenocepacia (strain MC0-3)	B1JTD4	8,00E-05	-	B
	Burkholderia multivorans (strain ATCC 17616 / 249)	A9AG59	8,00E-05	-	B
65	Dechloromonas aromatica (strain RCB)	Q47DF9	8,00E-05	-	B

ES 2 372 942 A1

	Ralstonia solanacearum (Pseudomonas solanacearum)	Q8XZR8	8,00E-05	-		B
5	Azoarcus sp. (strain EbN1)					
	(Aromatoleum aromaticum (strain EbN1))	Q5P172	1,00E-04	-		B
	Burkholderia dolosa AUO158	A2W9R2	1,00E-04	-		B
10	Cupriavidus taiwanensis R1 / LMG 19424 (Ralstonia taiwanensis)	B3R2N4	1,00E-04	-		B
	Janthinobacterium sp. (strain Marseille) (Minibacterium massiliensis)	A6SXX5	1,00E-04	-		B
15	Thiobacillus denitrificans (strain ATCC 25259)	Q3SKI5	1,00E-04	-		B
	Acidovorax avenae subsp. citrulli (strain AAC00-1)	A1TLJ1	2,00E-04	-		B
	Burkholderia ambifaria (strain MC40-6)	B1YRH5	2,00E-04	-		B
20	Burkholderia ambifaria MEX-5	B1T3P4	2,00E-04	-		B
	Marinobacter sp. ELB17	A3J9B4	2,00E-04	-		B
	Oceanobacter sp. RED65	Q1N0D7	2,00E-04	-		B
25	Ralstonia metallidurans (strain CH34 / ATCC 43123 / DSM 2839)	Q1LLY1	2,00E-04	-		B
	Burkholderia cenocepacia (strain AU 1054)	Q1BH16	3,00E-04	-		B
	Burkholderia cenocepacia (strain HI2424)	A0K7X8	3,00E-04	-		B
30	Burkholderia cenocepacia PC184	A2VXS8	3,00E-04	-		B
	Burkholderia cepacia J2315 / LMG 16656 (B. cenocepacia J2315)	B4EB12	3,00E-04	-	-	B
	Methylococcus capsulatus	Q60B83	4,00E-04	-		B
35	Acidovorax sp. (strain JS42)	A1WAQ4	5,00E-04	-		B
	Burkholderia ambifaria ATCC BAA-244 / AMMD (B. cepacia AMMD)	Q0BES5	5,00E-04	-		B
	Burkholderia ambifaria IOP40-10	B1FNI3	5,00E-04	-		B
40	Paenibacillus sp. JDR-2	B1DDQ0	5,00E-04	-		B
	Pseudomonas fluorescens (strain Pf0-1)	Q3KHI6	5,00E-04	-		B
	Ralstonia pickettii (strain 12J)	B2UAT0	6,00E-04	-		B
	Ralstonia pickettii 12D	A7CD46	6,00E-04	-		B
45	Bordetella petrii (strain ATCC BAA-461 / DSM 12804 / CCUG 43448)	A9IQR6	8,00E-04	-		B
	Leptothrix cholodnii ATCC 51168 / LMG 8142 / SP-6 (L. discophora SP-6)	B1Y1J3	8,00E-04	-		B
50	Marinomonas sp. (strain MWYL1)	A6W1T7	8,00E-04	-		B
	Neisseria gonorrhoeae (strain NCCP11945)	B4RLT7	8,00E-04	-		B
55	Polynucleobacter necessarius (strain STIR1)	B1XTP0	8,00E-04	-		B
	Pseudomonas fluorescens (strain Pf-5 / ATCC BAA-477)	Q4KHP9	8,00E-04	-		B
60	Ralstonia solanacearum (Pseudomonas solanacearum)	B5S1J7	8,00E-04	-		B
	Ralstonia solanacearum IPO1609	B5SLM8	8,00E-04	-		B
	Variovorax paradoxus S110	B7YY57	8,00E-04	-		B
	Variovorax paradoxus S110	B7YHV6	8,00E-04	-		B
65	Marinobacter algicola DG893	A6F295	0,001	-		B
	Nitrosomonas eutropha (strain C91)	Q0AHQ4	0,001	-		B

ES 2 372 942 A1

	Nodularia spumigena CCY 9414	A0ZEU3	0,001	-	B
5	Polaromonas naphthalenivorans (strain CJ2)	A1VPG1	0,001	-	B
	Polaromonas sp. (strain JS666 / ATCC BAA-500)	Q12BH5	0,001	-	B
	Pseudomonas aeruginosa	P29365	0,001	-	B
10	Pseudomonas aeruginosa (strain UCBPP-PA14)	Q02RK9	0,001	-	B
	Pseudomonas aeruginosa 2192	A3LEI7	0,001	-	B
	Pseudomonas aeruginosa C3719	A3KY07	0,001	-	B
	Pseudomonas aeruginosa LESB58	B7UZM6	0,001	-	B
15	Synechococcus sp. PCC 7335	B4WH29	0,001	-	B

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

- 5 1. Enzima con actividad 1-desoxi-D-xilulosa 5-fosfato reductoisomerasa que cataliza la reacción de producción de 2-C-metil-D-eritritol 4-fosfato (MEP) a partir de 1-desoxi-D-xilulosa 5-fosfato (DXP), que tiene la secuencia de aminoácidos SEQ ID NO: 41.
2. Enzima según la reivindicación 1, que es de *Brucella abortus*.
- 10 3. Enzima según la reivindicación 2, que es de *Brucella abortus* 2308.
4. Vector de expresión que permite la expresión en bacterias de la enzima definida en cualquiera de las reivindicaciones 1-3 activa, que comprende la secuencia de nucleótidos SEQ ID NO: 51.
- 15 5. Uso de una enzima según se define en cualquiera de las reivindicaciones 1-3, para la síntesis de isoprenoides.
6. Uso según la reivindicación 5, para la síntesis de MEP.
7. Método de obtención de MEP que comprende:
- 20 (a) expresar la enzima activa definida en cualquiera de las reivindicaciones 1-3 en un sistema celular; y
- (b) cultivar el sistema celular del paso (a) bajo condiciones y en un medio que permiten la producción y acumulación de MEP.
- 25 8. Método de obtención de MEP que comprende:
- (a) proporcionar un sistema celular que permita la expresión de la enzima definida en cualquiera de las reivindicaciones 1-3;
- 30 (b) cultivar el sistema celular del paso (a) bajo condiciones que permiten la producción de la enzima; y
- (c) cultivar el sistema celular bajo condiciones y en un medio que permiten la producción y acumulación de MEP o bien purificar el enzima y utilizarlo en un sistema *in vitro* bajo condiciones y en un medio que permiten la obtención de MEP, donde el medio comprende un sustrato seleccionado del grupo que consiste en DXP y una mezcla de piruvato y gliceraldehído 3-fosfato.

40

45

50

55

60

65

FIG. 1

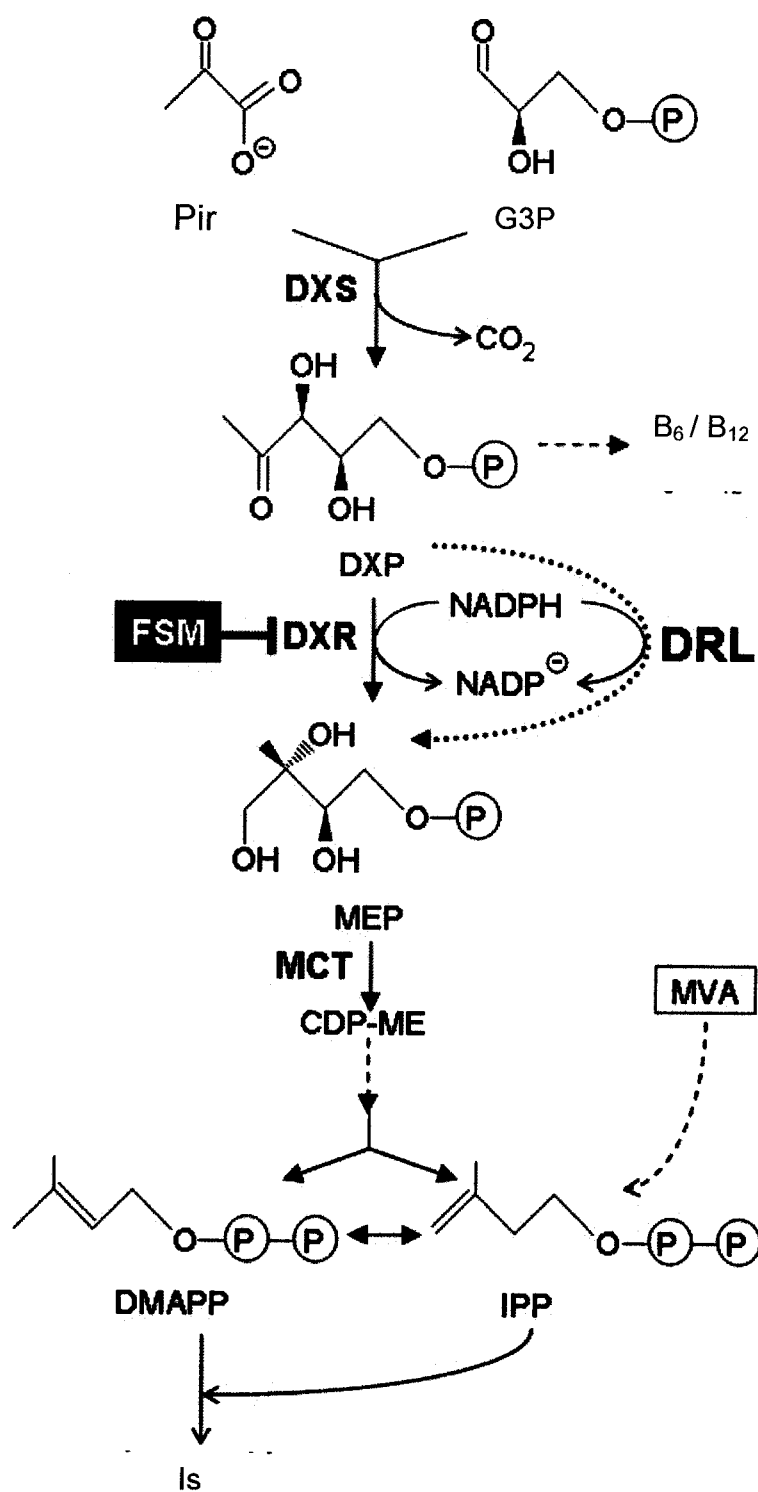


FIG. 2

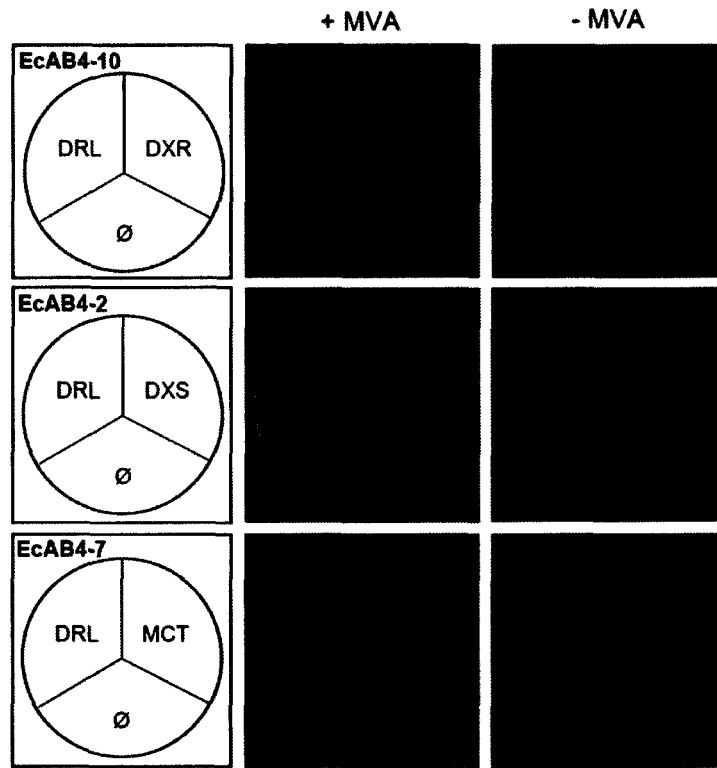


FIG. 3

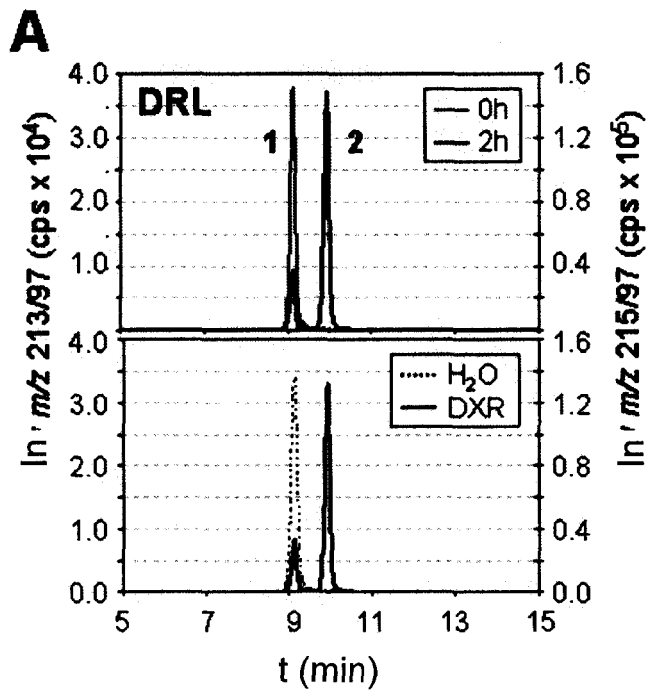


FIG. 3

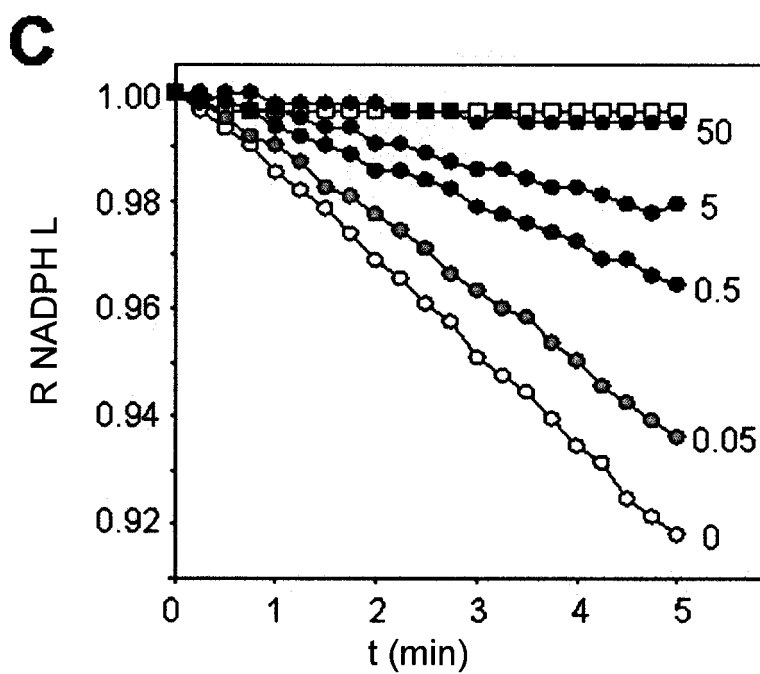
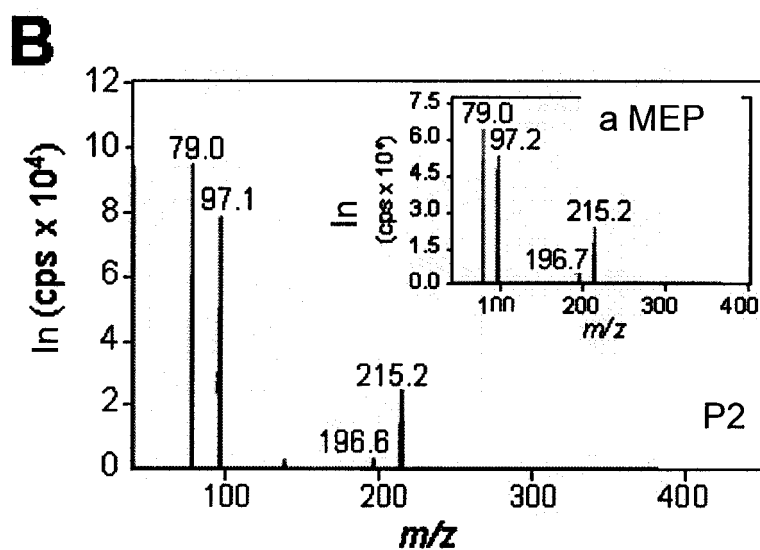


FIG. 4

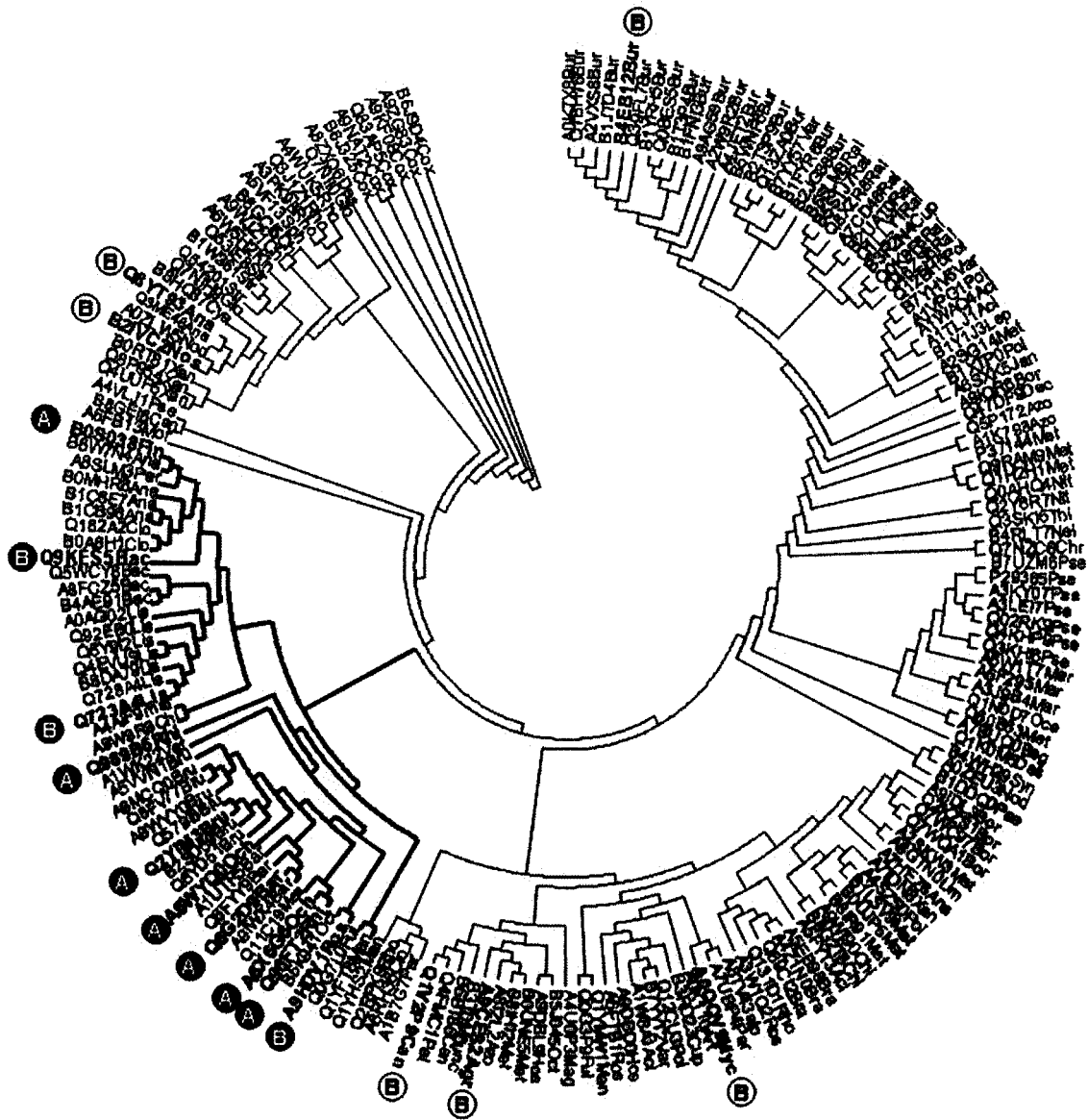


FIG. 5

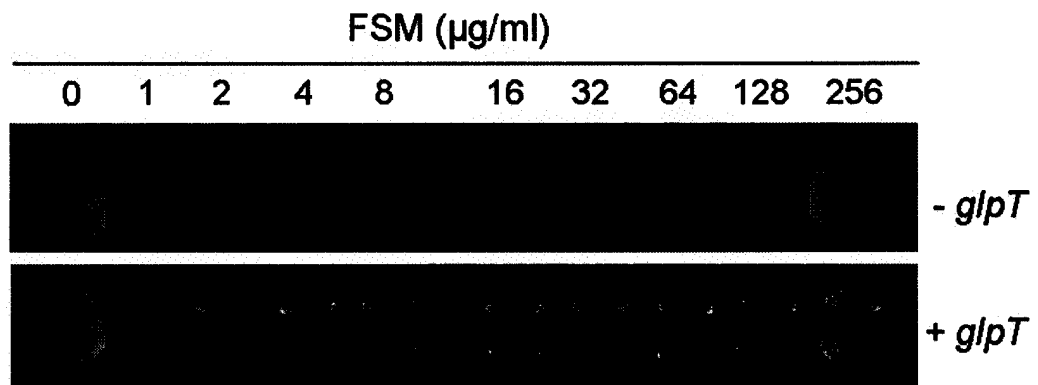


FIG. 6

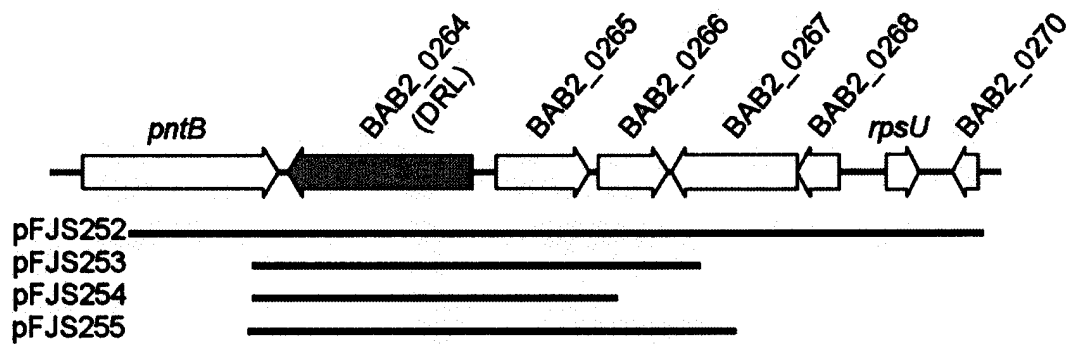


FIG. 7

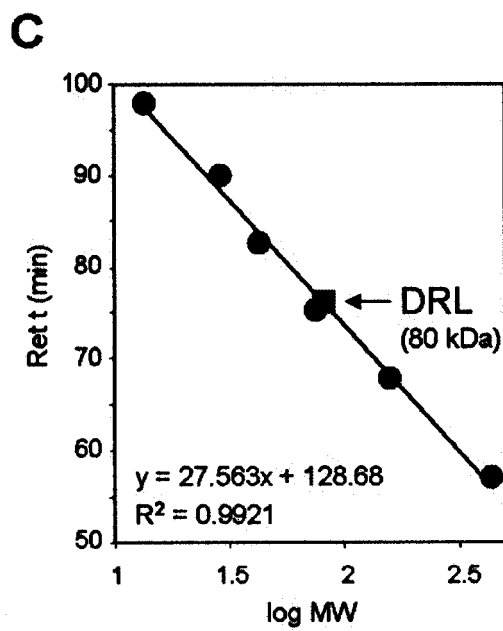
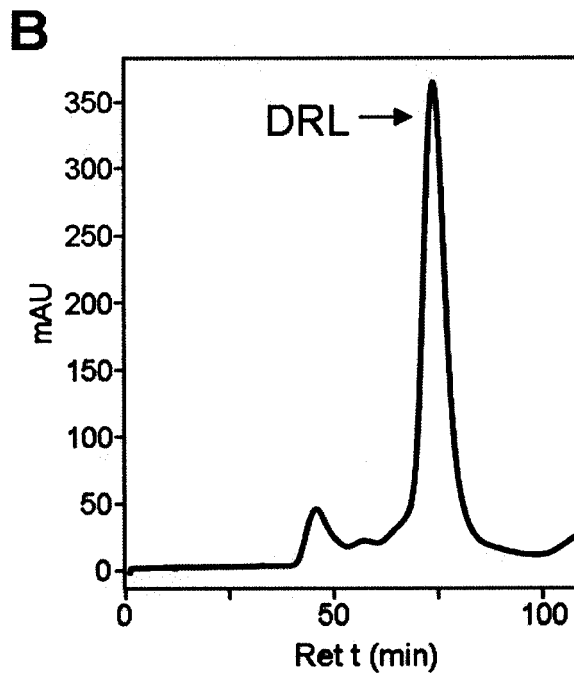
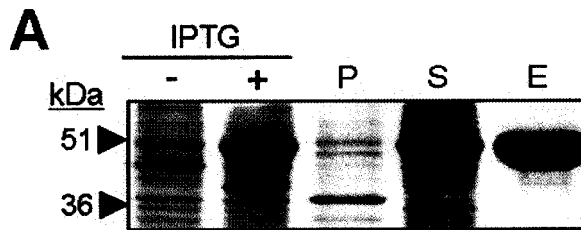


FIG. 8

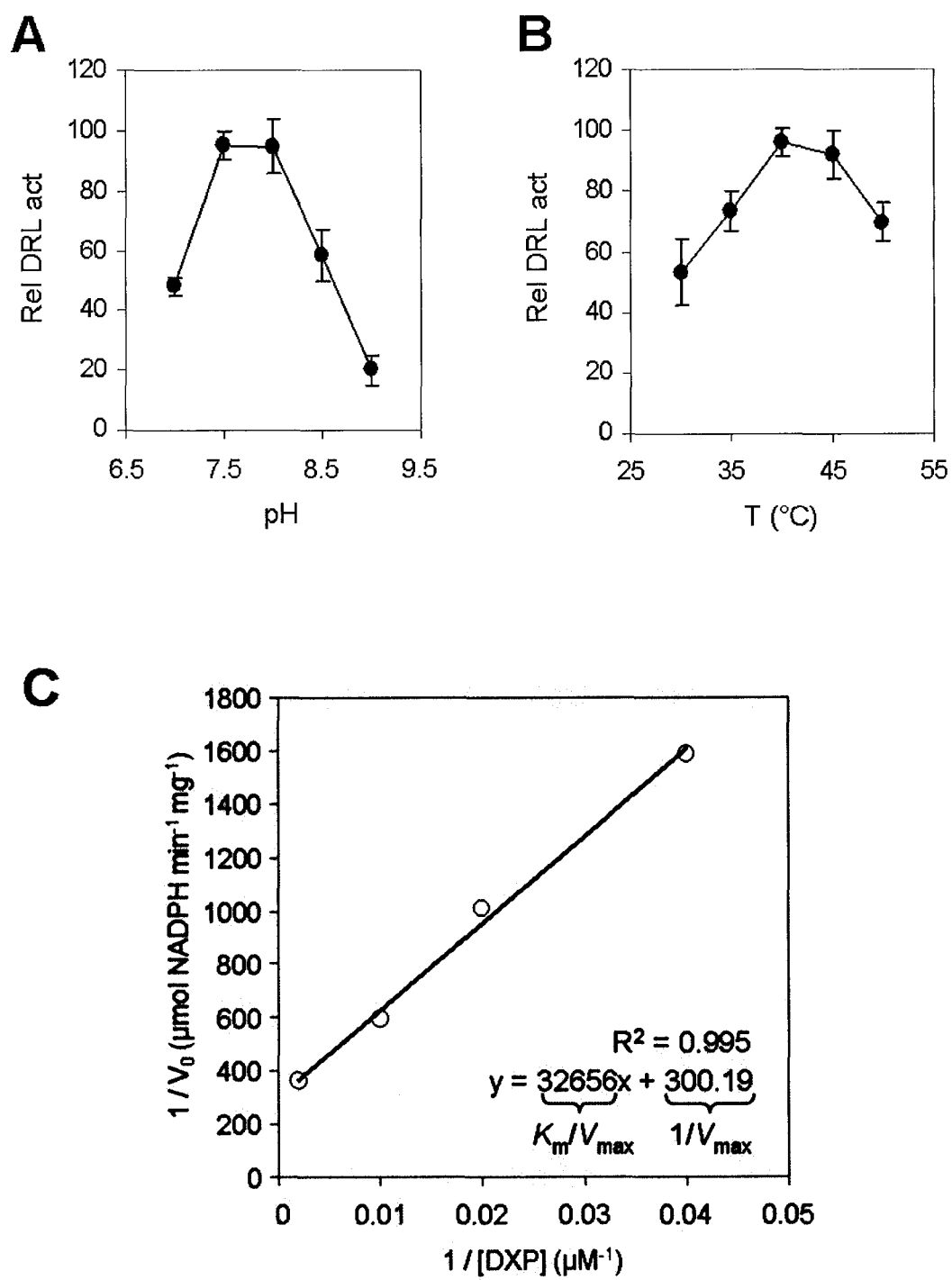
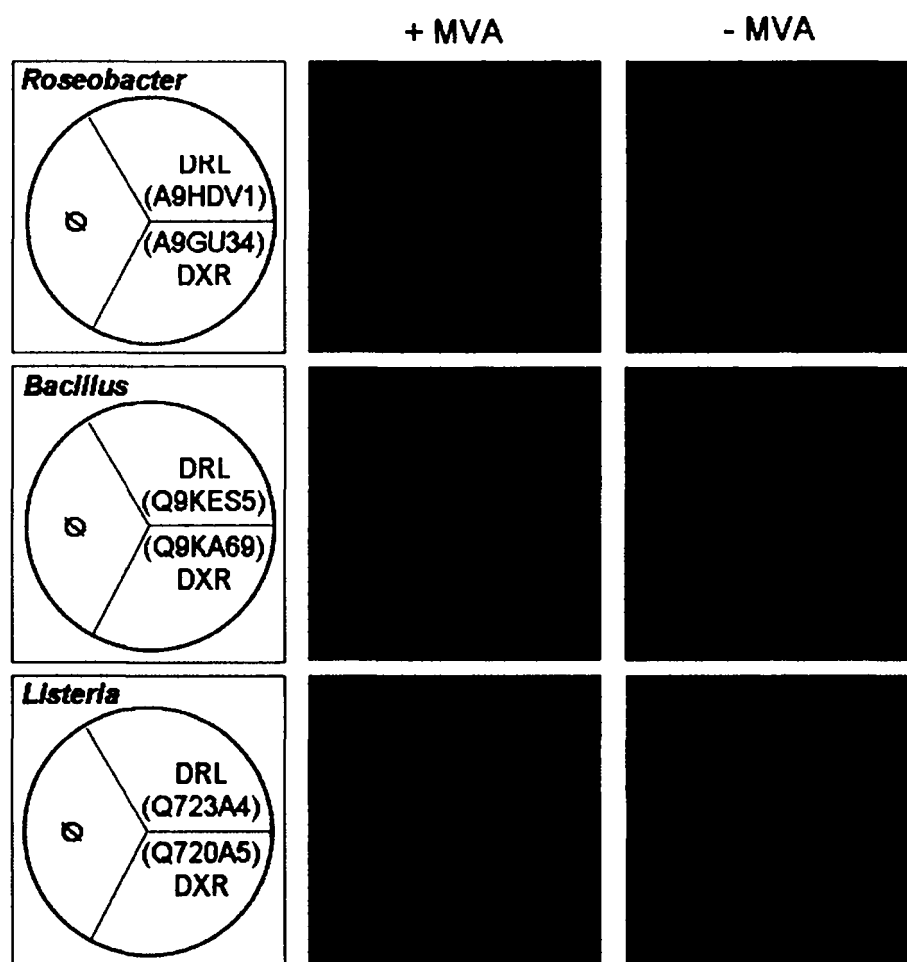


FIG. 9



ES 2 372 942 A1

LISTA DE SECUENCIAS

- <110> Universidad de Cantabria
Consejo Superior de Investigaciones Científicas
5 Universidad de Barcelona
Universidad de Almería
- <120> Nueva enzima para la biosíntesis de isoprenoides
- 10 <130> P1701ES00
- <160> 51
- 15 <170> PatentIn version 3.5
- <210> 1
<211> 30
20 <212> DNA
<213> Artificial
- 25 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano
- <400> 1
30 catggtcgac agaaacggca ggttctctca 30
- <210> 2
<211> 30
35 <212> DNA
<213> Artificial
- <220>
40 <223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano
- <400> 2
45 catgctgcag gtttttcagc gtcaatttca 30
- <210> 3
<211> 27
<212> DNA
50 <213> Artificial
- <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano
- 55 <400> 3
ccgcatatga caacaaatgt ggcaactg 27
- 60 <210> 4
<211> 29
<212> DNA
<213> Artificial Sequence
- 65 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

ES 2 372 942 A1

	<400> 4		
	atctcgagtt gccctaacat agcatcctg		29
5	<210> 5		
	<211> 20		
	<212> DNA		
	<213> Artificial Sequence		
10	<220>		
	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
15	<400> 5		
	cacgagatcg aaaggggata		20
	<210> 6		
20	<211> 20		
	<212> DNA		
	<213> Artificial Sequence		
25	<220>		
	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
	<400> 6		
30	tctggttctc gaacctcacc		20
	<210> 7		
	<211> 20		
35	<212> DNA		
	<213> Artificial Sequence		
	<220>		
40	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
	<400> 7		
45	tcgactgccg ttttaagtcc		20
	<210> 8		
	<211> 20		
50	<212> DNA		
	<213> Artificial Sequence		
	<220>		
55	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
	<400> 8		
60	gtgcttcctt cagctgttcc		20
	<210> 9		
	<211> 20		
	<212> DNA		
65	<213> Artificial Sequence		

ES 2 372 942 A1

<220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

5 <400> 9
tcgtccattt ccacgatttt 20

10 <210> 10
<211> 20
<212> DNA
<213> Artificial Sequence

15 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

20 <400> 10
gaagcggcgg ataagataca 20

<210> 11
25 <211> 20
<212> DNA
<213> Artificial Sequence

30 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

35 <400> 11
ttatccagcc caacgaaaag 20

<210> 12
40 <211> 20
<212> DNA
<213> Artificial Sequence

45 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

50 <400> 12
gtcgatcgag ttgggaaagt 20

<210> 13
55 <211> 20
<212> DNA
<213> Artificial Sequence

60 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

65 <400> 13
ccaccgcat aatcctattt 20

ES 2 372 942 A1

<210> 14
<211> 20
<212> DNA
5 <213> Artificial Sequence

<220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano
10
<400> 14
gagggatcaa atggcaaaga 20

15 <210> 15
<211> 20
<212> DNA
<213> Artificial Sequence
20

<220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano
25
<400> 15
agcagaggtg cttttcagga 20

30 <210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

35 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano
40
<400> 16
acttgcgat caggaacatc 20

45 <210> 17
<211> 20
<212> DNA
<213> Artificial Sequence

50 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano
55
<400> 17
gacaaggcca gaatgaaagc 20

60 <210> 18
<211> 20
<212> DNA
<213> Artificial Sequence

65 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

ES 2 372 942 A1

	<400> 18		
	ccaattgatc ggggtatctg		20
5	<210> 19		
	<211> 20		
	<212> DNA		
	<213> Artificial Sequence		
10	<220>		
	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
15	<400> 19		
	ctttcaatgg cgaaacgaat		20
	<210> 20		
20	<211> 20		
	<212> DNA		
	<213> Artificial Sequence		
25	<220>		
	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
30	<400> 20		
	tcattctctta tgcccgtct		20
	<210> 21		
35	<211> 20		
	<212> DNA		
	<213> Artificial Sequence		
40	<220>		
	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
45	<400> 21		
	caaaactcgcg acacaagatg		20
	<210> 22		
	<211> 20		
50	<212> DNA		
	<213> Artificial Sequence		
	<220>		
55	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
	<400> 22		
60	tcatgcccac cttgcattta		20
	<210> 23		
	<211> 20		
	<212> DNA		
65	<213> Artificial Sequence		

ES 2 372 942 A1

<220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

5 <400> 23
tgccggattg atagatgtga 20

<210> 24
10 <211> 20
<212> DNA
<213> Artificial Sequence

15 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

20 <400> 24
cctccttaga gggtaggaag 20

<210> 25
<211> 20
<212> DNA
25 <213> Artificial Sequence

<220>
30 <223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

<400> 25
aagcaaaaat tttcgggtcca 20

35 <210> 26
<211> 20
<212> DNA
40 <213> Artificial Sequence

<220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

45 <400> 26
taccagggca tggtaggtatt 20

<210> 27
50 <211> 20
<212> DNA
<213> Artificial Sequence

55 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

60 <400> 27
ttcatcggcg tattcttgtg 20

<210> 28
65 <211> 22
<212> DNA
<213> Artificial Sequence

ES 2 372 942 A1

<220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

5 <400> 28
aaaatcagcc gaaatttcaa ag 22

<210> 29
10 <211> 20
<212> DNA
<213> Artificial Sequence

15 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

20 <400> 29
caccctaca gccaattcat 20

<210> 30
25 <211> 20
<212> DNA
<213> Artificial Sequence

30 <220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

35 <400> 30
taatgtctcg tcagggatgg 20

<210> 31
<211> 20
40 <212> DNA
<213> Artificial Sequence

<220>
45 <223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

50 <400> 31
aactatcccc atcggtagcc 20

<210> 32
<211> 20
<212> DNA
55 <213> Artificial Sequence

<220>
<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

60 <400> 32
actgacgcac cctaccaaac 20

<210> 33
65 <211> 19
<212> DNA
<213> Artificial Sequence

ES 2 372 942 A1

	<220>		
	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
5	<400> 33		
	gggaacgcat tctcctcag		19
	<210> 34		
10	<211> 20		
	<212> DNA		
	<213> Artificial Sequence		
15	<220>		
	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
	<400> 34		
20	ccattggaca tgtcagcttc		20
	<210> 35		
	<211> 20		
25	<212> DNA		
	<213> Artificial Sequence		
30	<220>		
	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
	<400> 35		
35	atgacggtat gcacccaaat		20
	<210> 36		
	<211> 20		
40	<212> DNA		
	<213> Artificial Sequence		
45	<220>		
	<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano		
	<400> 36		
	ggatggactg gccaatgtta		20
50	<210> 37		
	<211> 20		
	<212> DNA		
55	<213> Artificial Sequence		
	<400> 37		
	atagatcgct gccaatccac		20
60	<210> 38		
	<211> 20		
	<212> DNA		
65	<213> Artificial Sequence		

ES 2 372 942 A1

<220>

<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

5 <400> 38

gttcgacggt ctgaccaact

20

10

<210> 39

<211> 19

<212> DNA

15 <213> Artificial Sequence

<220>

<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

20

<400> 39

gaacctcgac gacctgacc

19

25

<210> 40

<211> 20

<212> DNA

30

<213> Artificial Sequence

<220>

<223> oligonucleótido para amplificación por PCR de ADN genómico bacteriano

35

<400> 40

ggatgtcggg gaacgtatgg

20

40

<210> 41

<211> 438

45

<212> PRT

<213> *Brucella abortus*

<400> 41

50

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

55

Arg Ala Glu Thr Gly Lys Pro Ile Arg Ile Gly Leu Ile Gly Ala Gly
 20 25 30

60

65

ES 2 372 942 A1

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

ES 2 372 942 A1

Ser Lys Ser Gly Val Val Asp Tyr Ser Ile Gly Lys Gly Val Ser Pro
 275 280 285

5

Gly Val Phe Val Val Ala Lys Met Asp His Pro Arg Leu Asn Glu Arg
 290 295 300

10

Leu Glu Asp Leu Lys Ile Gly Lys Gly Pro Tyr Phe Thr Phe His Arg
 305 310 315 320

15

Pro Tyr His Leu Thr Ser Leu Glu Val Pro Leu Thr Val Ala Arg Val
 325 330 335

20

Val Leu His Gly Lys Thr Asp Met Val Pro Leu Pro Lys Pro Val Ala
 340 345 350

25

Glu Val Cys Ala Val Ala Lys Lys Asp Met Gln Pro Gly Glu His Leu
 355 360 365

30

Asp Ala Ile Gly Gln Tyr Cys Tyr Arg Ser Trp Ile Met Thr Val Pro
 370 375 380

35

Glu Ala Arg Ala Ala Lys Ala Ile Pro Cys Gly Leu Leu Gln Asn Gly
 385 390 395 400

40

Thr Val Ile Ala Pro Ile Lys Lys Gly Glu Leu Ile Thr Tyr Ala Asn
 405 410 415

45

Ala Ala Pro Gln Pro Gly Ser Arg Ile Ala Glu Leu Arg Ala Leu Gln
 420 425 430

45

Asp Ala Met Leu Gly Gln
 435

50

<210> 42
 <211> 438
 <212> PRT
 <213> *Ochrobactrum anthropi*

55

<400> 42

60

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

65

Arg Ala Glu Thr Gly Lys Pro Ile Arg Ile Gly Leu Ile Gly Ala Gly
 20 25 30

65

Glu Met Gly Thr Asp Ile Val Thr Gln Val Ala Arg Met Gln Gly Ile

ES 2 372 942 A1

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

ES 2 372 942 A1

Gly Val Phe Val Ile Ala Lys Met Glu His Pro Arg Leu Val Glu Arg
 290 295 300
 5

Leu Glu Asp Leu Lys Met Gly Lys Gly Pro Tyr Phe Thr Phe His Arg
 305 310 315 320
 10

Pro Phe His Leu Thr Ser Leu Glu Val Pro Leu Thr Val Ala Arg Val
 325 330 335
 15

Val Leu His Gly Lys Thr Asp Met Val Pro Leu Pro Lys Pro Val Ala
 340 345 350
 20

Glu Val Cys Ala Val Ala Lys Lys Asp Leu Gln Pro Gly Asp His Leu
 355 360 365
 25

Asp Ala Ile Gly Gln Tyr Cys Tyr Arg Ser Trp Ile Met Thr Thr Pro
 370 375 380
 30

Glu Ala His Ala Ala Lys Ala Ile Pro Cys Gly Leu Leu Gln Asn Gly
 385 390 395 400
 35

Thr Val Ile Ala Pro Ile Arg Lys Gly Glu Leu Ile Thr Tyr Ala Asn
 405 410 415
 40

Ala Ala Pro Gln Pro Gly Ser Lys Ile Ala Glu Leu Arg Ala Leu Gln
 420 425 430
 45

Asp Lys Met Ile Tyr Gly
 435

<210> 43
 <211> 439
 <212> PRT
 50 <213> *Ochrobactrum anthropi*
 <400> 43

Met Ala Thr Asn Val Ala Leu Thr Gly Leu Ala Arg Asp Met Gln Ala
 1 5 10 15
 55

Arg Ala Asp Ser Gly Arg Pro Ile Arg Ile Gly Leu Ile Gly Ser Gly
 20 25 30
 60

Glu Met Gly Thr Asp Ile Val Thr Arg Val Ala His Met Pro Gly Ile
 35 40 45
 65

ES 2 372 942 A1

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

5
 Val Asp Ile Ala Phe Gln Glu Glu Gly His Gly Arg Glu Val Ser Thr
 65 70 75 80

10
 Ala Ser Asp Leu Thr Ala Ala Met Glu Ala His Lys Val Ala Val Thr
 85 90 95

15
 Asp Asn Ala Asn Leu Ile Leu Glu Asn Asp Leu Ile Asp Val Val Ile
 100 105 110

20
 Asp Ala Thr Gly Val Pro Ala Val Gly Ala Glu Ile Gly Leu Arg Ala
 115 120 125

25
 Met Glu Tyr Gly Lys His Leu Val Met Met Asn Val Glu Ala Asp Val
 130 135 140

30
 Thr Ile Gly Ala Tyr Leu Lys Ala Glu Ala Glu Arg Leu Gly Val Thr
 145 150 155 160

35
 Tyr Ser Leu Gly Ala Gly Asp Glu Pro Ser Ser Cys Met Glu Leu Ile
 165 170 175

40
 Glu Phe Val Ser Ala Met Gly His Pro Ile Val Ala Ala Gly Lys Gly
 180 185 190

45
 Lys Asn Asn Pro Leu Asn Ile Asp Ala Val Pro Asp Gln Tyr Leu Glu
 195 200 205

50
 Glu Ala Thr Arg Arg Asn Met Asn Val Arg Met Leu Val Glu Phe Val
 210 215 220

55
 Asp Gly Ser Lys Thr Met Val Glu Met Ala Ala Ile Ala Asn Ala Thr
 225 230 235 240

60
 Gly Leu Val Pro Asp Lys Ala Gly Met His Gly Pro Ala Ala Thr Leu
 245 250 255

65
 Asp Gln Leu Asn Lys Thr Leu Ile Pro Glu Lys Asp Gly Gly Val Leu
 260 265 270

Ser Lys Val Gly Val Val Asp Tyr Ser Ile Gly Lys Gly Val Ala Pro
 275 280 285

Gly Val Phe Val Val Ala Asp Met Ser His Pro Arg Ile Ser Glu Arg

ES 2 372 942 A1

	290		295		300											
5	Met 305	Glu	Asp	Leu	Lys	Met 310	Gly	Lys	Gly	Pro	Tyr 315	Phe	Thr	Phe	His	Arg 320
10	Pro	Tyr	His	Leu	Thr 325	Ser	Leu	Glu	Val 330	Pro	Leu	Thr	Cys	Ala	Arg 335	Val
15	Val	Leu	Tyr	Gly 340	Lys	Pro	Asp	Met	Val 345	Pro	Leu	Ser	Lys	Pro	Val	Ala 350
20	Glu	Val	Ala 355	Ala	Val	Ala	Lys	Lys 360	Asp	Met	Gln	Pro	Gly 365	Glu	Lys	Leu
25	Asp	Ala	Ile	Gly	Glu	Tyr	Cys 375	Tyr	Arg	Ala	Trp	Ile	Met 380	Thr	Ser	Gly
30	Glu	Ala	Arg	Asp	Ala	His 390	Ala	Ile	Pro	Cys 395	Gly	Leu	Leu	Gln	Gly	Gly 400
35	Ser	Val	Thr	Lys	Pro	Ile	Lys 405	Lys	Gly	Glu	Leu	Ile	Thr	Tyr	Asp 415	Asn
40	Ala	Ala	Val	Ala 420	Pro	Gly	Ser	Lys	Ile 425	Ala	Glu	Leu	Arg	Ala	Arg 430	Gln
45	Asp	Lys	Leu	Val	Tyr	Gly	Ala									
50																
55	Met 1	Ala	Ser	Asn 5	Ile	Ser	Leu	Thr	Gly 10	Leu	Ala	Arg	Asp	Leu 15	Asp	Glu
60	Arg	Gly	Lys	Ser 20	Gly	Lys	Pro	Ile	Arg 25	Ile	Gly	Leu	Ile	Gly 30	Ser	Gly
65	Glu	Met	Gly 35	Thr	Asp	Ile	Val	Thr 40	Arg	Val	Ala	His	Met 45	Ser	Gly	Ile
70	Glu	Ile	Gly	Ala	Ile	Ser	Glu	Leu	Asn	Leu	Pro	Ala	Ala	Ser	Arg	Ala
75																

<210> 44
 <211> 442
 <212> PRT
 <213> *Mesorhizobium loti*
 <400> 44

ES 2 372 942 A1

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

ES 2 372 942 A1

	Met	Glu	Asp	Leu	Lys	Met	Gly	Lys	Gly	Pro	Tyr	Phe	Thr	Phe	His	Arg
	305					310					315					320
5	Pro	Tyr	His	Leu	Thr	Ser	Leu	Glu	Val	Pro	Leu	Thr	Cys	Ala	Arg	Val
					325					330					335	
10	Val	Leu	Tyr	Gly	Lys	Ala	Asp	Met	Val	Pro	Leu	Ala	Lys	Pro	Val	Ala
				340					345					350		
15	Glu	Val	Ala	Ala	Val	Ala	Lys	Lys	Asp	Met	Gln	Pro	Gly	Glu	Lys	Leu
			355					360					365			
20	Asp	Ala	Ile	Gly	Glu	Tyr	Cys	Tyr	Arg	Ala	Trp	Ile	Met	Thr	Ala	Pro
	370						375					380				
25	Glu	Ala	His	Ala	Ala	Lys	Ala	Ile	Pro	Cys	Gly	Leu	Leu	Gln	Gly	Gly
	385					390					395					400
30	Ser	Val	Thr	Ala	Pro	Ile	Lys	Lys	Gly	Glu	Leu	Ile	Thr	Tyr	Ala	Asn
					405					410					415	
35	Ala	Val	Pro	Ala	Pro	Gly	Ser	Lys	Ile	Ala	Glu	Leu	Arg	Ala	Arg	Gln
				420					425					430		
40	Asp	Lys	Leu	Val	Tyr	Gly	Thr	Val	Gly	Ala						
		435						440								
45	<210> 45															
	<211> 433															
50	<212> PRT															
	<213> <i>Mesorhizobium loti</i>															
55	<400> 45															
60	Met	Thr	Gly	Leu	Ala	Arg	Asp	Leu	Ala	Ser	Arg	Ala	Ala	Glu	Gly	Arg
	1				5					10				15		
65	Pro	Val	Arg	Ile	Gly	Val	Ile	Gly	Ser	Gly	Glu	Met	Gly	Thr	Asp	Leu
				20				25						30		
70	Val	Thr	Gln	Gly	Met	Leu	Met	Pro	Gly	Ile	Ser	Val	Cys	Ala	Val	Ser
		35						40					45			
75	Thr	Arg	Arg	Pro	His	Thr	Ala	Arg	Asp	Ala	Ile	Arg	Ile	Ala	Tyr	Gly
	50						55					60				
80	Asp	Glu	Ala	Met	Ala	Val	Glu	Ala	Asp	Ala	Ala	Ser	Lys	Val	Thr	Ala
	65					70					75					80

ES 2 372 942 A1

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

ES 2 372 942 A1

Glu Val Pro Leu Thr Ala Ala Arg Ile Val Leu Phe Gly Lys Pro Asp
 325 330 335
 5
 Met Val Pro Leu Pro Arg Pro Val Ala Glu Val Cys Ala Val Ala Lys
 340 345 350
 10
 Arg Asp Leu Ala Ala Gly Glu Thr Phe Asp Ala Ile Gly Glu Thr Cys
 355 360 365
 15
 Tyr Arg Ser Trp Thr Met Thr Val Gly Glu Ala Arg Ala Gln His Ala
 370 375 380
 20
 Val Pro Val Gly Leu Leu Glu Gly Gly Lys Val Leu Lys Pro Val Arg
 385 390 395 400
 25
 Lys Gly Glu Leu Leu Thr Ala Asp Asn Ala Ala Pro Asp Gln Thr Thr
 405 410 415
 30
 Arg Leu Phe Ala Leu Arg Arg Leu Gln Asp Glu Met Leu Tyr Gly Ala
 420 425 430
 35
 Asn
 <210> 46
 <211> 445
 <212> PRT
 <213> *Bartonella henselae*
 40
 <400> 46
 45
 Met Ala Ser Asn Val Ser Leu Thr Gly Leu Ala His Asp Leu Lys Gln
 1 5 10 15
 50
 Arg Ala Glu Asn His Pro Pro Ile Arg Ile Gly Leu Ile Gly Cys Gly
 20 25 30
 55
 Glu Met Gly Thr Asp Leu Leu Ser Ser Ile Ala His Met Asp Gly Ile
 35 40 45
 60
 Thr Val Ala Ala Val Ala Thr Arg Thr Pro Ser Arg Ile Phe Asp Ala
 50 55 60
 65
 Ala Arg Leu Ala Tyr Gly Glu Glu Gly His Val Arg Glu Val Glu Asn
 65 70 75 80
 65
 Ala Ser Ala Leu Thr Arg Thr Ile Glu Lys Gly Leu Ile Ala Ala Thr

ES 2 372 942 A1

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

ES 2 372 942 A1

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

5

Glu Val Cys Ala Val Ala Lys Lys Asp Leu Tyr Pro Gly Asp Gln Leu
 355 360 365

10

Asp Phe Ile Gly Leu Tyr Ser Tyr Arg Ala Trp Ile Met Asn Ile Ala
 370 375 380

15

Glu Ala Arg Met His Gln Ala Ile Pro Cys Gly Leu Leu Glu Asn Ala
 385 390 395 400

20

Thr Val Thr Ala Glu Ile Lys Lys Asn Glu Leu Ile Thr Val His Asn
 405 410 415

25

Thr Ala Ile Arg Glu Asp Gln Trp Ile Ala Arg Leu Arg Thr Lys Gln
 420 425 430

30

Asp Leu Leu Leu Asn Ala Ser Ser Leu Pro Ser His Ala
 435 440 445

<210> 47

<211> 443

35 <212> PRT

<213> *Roseobacter litoralis*

<400> 47

40

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

45

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

50

Glu Met Gly Thr Asp Ile Val Thr Arg Ala Gly Met Met Asp Gly Val
 35 40 45

55

Glu Val Ala Ala Ile Ser Glu Val Asn Pro Ala Ala Ala His Lys Ala
 50 55 60

60

Val Glu Ile Ala His Ala Ala Ala Gly Phe Ser Ala Asp Ala Asn Ser
 65 70 75 80

65

Thr Asp Ala Leu Asn Ala Ala Ile Glu Asn Gly Lys Thr Ala Val Ala
 85 90 95

ES 2 372 942 A1

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

ES 2 372 942 A1

	340	345	350
5	Glu Val Ala Ala Val Ala Lys Arg Asp Leu Lys Ala Gly Glu Thr Leu 355 360 365		
10	Asp Gln Ile Gly Glu Tyr Thr Tyr Arg Ala Trp Ala Met Glu Thr Ser 370 375 380		
15	Arg Ala Arg Thr Ala Arg Ala Leu Pro Ala Gly Leu Leu Thr Gly Ala 385 390 395 400		
20	Val Thr Thr Ala Ala Ile Ala Lys Gly Glu Leu Ile Thr Ala His Asn 405 410 415		
25	Thr Thr Leu Pro Asp Ala Arg Ile Val Glu Leu Arg Arg Arg Gln Asp 420 425 430		
30	Glu Met Ile Tyr Gly Lys Asp Ala Ala Asn Val 435 440		
35	<210> 48 <211> 439 <212> PRT <213> <i>Bacillus halodurans</i> <400> 48		
40	Met Leu Gly Ile Asn Arg Lys Leu Glu Gln Leu Glu Ser Asp Gly Thr 1 5 10 15		
45	Ile Ile Arg Val Gly Leu Val Gly Ala Gly Gln Met Gly Arg Gly Met 20 25 30		
50	Val Ser Gln Ile Glu Ser Met Lys Gly Met Arg Val Val Val Thr Ala 35 40 45		
55	Asp Ile Val Leu Glu Asn Val Lys Leu Ala Tyr Glu Arg Ala Gly Val 50 55 60		
60	Pro Lys Asp Glu Val Ile Glu Thr Asp Glu Ile Ser Lys Ala Ala Asp 65 70 75 80		
65	Ala Ile Gln Ala Gly Lys Val Val Ala Thr Gly Asp Ala Glu Leu Val 85 90 95		
70	Thr Ala Leu Ser Glu Val Asp Val Val Val Asp Ala Thr Gly Ile Pro 100 105 110		

ES 2 372 942 A1

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

ES 2 372 942 A1

Asp Leu Ala Ala Gly Glu Phe Leu Asp Ser Ile Gly Gly Phe Thr Val
 355 360 365

5

Tyr Gly Thr Ile Leu Thr Val Thr Asp Ala Lys Glu Lys Gly Ala Leu
 370 375 380

10

Pro Ile Gly Leu Val Asp Ala His Val Gln Val Lys Arg Pro Ile Lys
 385 390 395 400

15

Lys Gly Glu Ile Ile Thr Tyr Asp Asp Val Glu Gln Lys Lys Glu Ser
 405 410 415

20

Thr Ile Trp Arg Leu Arg Arg Ile Gln Asp Glu Thr Leu Val Gly Lys
 420 425 430

25

Arg Glu Leu Glu Ala Lys Val
 435

<210> 49

<211> 416

30 <212> PRT

<213> *Listeria monocytogenes*

<400> 49

35

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

40

Ile Arg Val Gly Val Ile Gly Ala Gly Gln Met Gly Phe Gly Met Ile
 20 25 30

45

Ser Gln Ile Ala Ala Ile Pro Gly Met Ser Ile Val Gly Ile Ser Asp
 35 40 45

50

Ile His Val Glu Ala Ala Gln Lys Ala Ala Asp Ala Tyr Asn Ala Thr
 50 55 60

55

Ala Thr Lys Lys Glu Lys Ile Leu Leu Ser Asn Asp Phe Lys Glu Ile
 65 70 75 80

60

Ile His Ser Asp Leu Val Glu Val Ile Val Asp Ala Thr Gly Val Pro
 85 90

65

Glu Val Gly Ala Lys Ile Ser Leu Glu Thr Leu Leu Ala Lys Lys Gln
 100 105 110

Leu Val Leu Leu Asn Val Glu Ile Asp Ile Thr Ile Gly Pro Leu Met
 115 120 125

ES 2 372 942 A1

Lys Lys Leu Tyr Asp Ser Ala Gly Leu Val Tyr Thr Gly Ser Asp Gly
 130 135 140
 5

Asp Glu Pro Ala Ala Ile Thr Glu Leu Tyr Glu Phe Ser Lys Ser Met
 145 150 155 160
 10

Gly Met Glu Val Leu Val Ala Gly Lys Gly Lys Asn Asn Lys Leu Lys
 165 170 175
 15

Ile Ser Ala Asn Pro Asp Ser Cys Gln Ala Glu Ala Asp Gly Lys Asn
 180 185 190
 20

Met Ala Ser His Met Leu Ala Ala Phe Gln Asp Gly Thr Lys Thr Met
 195 200 205
 25

Ala Glu Met Asn Leu Leu Ser Asn Ala Ile Gly Tyr Val Pro Asp Val
 210 215 220
 30

Val Gly Met His Gly Ile Ser Gly Asp Val Asp Ser Val Ile Lys Asp
 225 230 235 240
 35

Leu Asp Leu Lys Asp Gln Gly Gly Ile Leu Asn Lys Phe Gly Val Val
 245 250 255
 40

Glu Tyr Val Asp Gly Leu Ala Pro Gly Val Phe Val Ile Val Lys Gly
 260 265 270
 45

Gln Asn Glu Gly Val Ser His Glu Leu Ser Tyr Leu Met Lys Lys Gly
 275 280 285
 50

Asp Arg Asp His His Ile Leu Tyr Arg Pro Tyr His Leu Ala Ser Leu
 290 295 300
 55

Glu Thr Pro Leu Thr Ile Ala Lys Ala Val Leu Asn His Asp His Ala
 305 310 315 320
 60

Ile Val Pro Met Gly Ala Pro Val Ser Glu Thr Val Ala Val Ala Lys
 325 330 335
 65

Lys Asp Ile Ala Ala Gly Glu Lys Leu Asp Gly Ile Gly Gly Phe Cys
 340 345 350
 70

Val Arg Gly Val Leu Glu Thr His Val Asp Met Ala Thr Asn Gly His
 355 360 365
 75

ES 2 372 942 A1

Ile Pro Ile Gly Leu Ile Ser Gly Glu Val Val Ala Arg Arg Asn Ile
 370 375 380

5

Lys Ala Gly Thr Phe Ile Thr Asp Glu Asp Val Ser Leu Asp Glu Ser
 385 390 395 400

10

Thr Thr Val Trp Lys Leu Arg Lys Leu Gln Asp Glu Thr Phe Asn Lys
 405 410 415

15 <210> 50
 <211> 430
 <212> PRT
 <213> *Finegoldia magna*

20 <400> 50

25 Met Asn Lys Met Gln Arg Lys Leu Val Asp Leu Asp Asn His Gly Gln
 1 5 10 15

30 Lys Val Arg Val Gly Leu Val Gly Cys Gly Lys Met Gly Ser Gly Leu
 20 25 30

35 Val Ser Gln Leu Ser Arg Ile Lys Gly Met Arg Pro Ser Val Ile Ile
 35 40 45

40 Asp Arg His Val Asp Lys Cys Val Thr Ala Leu Arg Lys Ala Gly Val
 50 55 60

45 Lys Asp Val Asp Ile Val Arg Thr Thr Asp Leu Lys Val Ala Glu Asn
 65 70 75 80

50 Ala Ile Lys Asn Asn Ser Phe Val Val Ser Asp Asp Tyr Thr Leu Ser
 85 90 95

55 Tyr Lys Leu Asp Met Ile Asp Gly Val Ile Asp Ala Thr Gly Asn Pro
 100 105 110

60 Pro Phe Gly Thr Gln Leu Ala Val Glu Ser Ile Glu His Glu Lys His
 115 120 125

65 Thr Ile Leu Leu Asn Val Glu Cys Asp Ala Val Val Gly Pro Ile Leu
 130 135 140

Asn Glu Met Ala Lys Lys Lys Gly Val Val Tyr Thr Gly Ser Ala Gly
 145 150 155 160

65 Asp Glu Pro Gly Ala Ile Ile Gln Leu Ser Asp Phe Ala Leu Gly Leu

ES 2 372 942 A1

	165	170	175
5	Gly Phe Lys Leu Leu Ala Val Gly Lys Gly Lys Asn Asn Pro Leu Asp 180	185	190
10	Asn Tyr Thr Asn Glu Asp Ile Leu Arg Glu Glu Ala Leu Ser Lys Gly 195	200	205
15	Leu Val Pro Lys Met Leu Thr Ser Phe Val Asp Gly Thr Asn Thr Met 210	215	220
20	Ile Glu Leu Thr Ala Val Ala Asn Ala Leu Gly Phe Thr Pro Asp Val 225	230	235
25	Leu Gly Cys His Gly Ile Thr Thr Asn Ile His Asp Ile Ala Asp Lys 245	250	255
30	Phe Lys Leu Lys Glu Gln Gly Gly Ile Leu Asn Asn Tyr Asn Ile Val 260	265	270
35	Asp Phe Ala Phe Gly Val Ala Pro Gly Val Phe Ala Ile Val Thr His 275	280	285
40	Asp Thr Asn Glu Val His Asp Leu Met Glu Tyr Leu Ser Met Gly Lys 290	295	300
45	Gly Pro Asn Tyr Thr Leu Tyr Arg Pro Tyr His Leu Thr Ser Leu Glu 305	310	315
50	Thr Pro Leu Thr Ile Tyr Asp Ala Ile Val Glu Lys Glu Pro Thr Ile 325	330	335
55	Tyr Pro Lys Tyr Gly Gln Val Ser Asp Thr Ile Thr Val Ala Lys Arg 340	345	350
60	Asp Ile Lys Lys Gly Gln Lys Ile Glu Gly Ile Gly Gly His Asp Cys 355	360	365
65	Phe Gly Lys Ile Thr Ser His Lys His Gln Met Glu Asn Asn Leu Leu 370	375	380
70	Pro Met Pro Ile Ile Thr Glu Lys Thr Thr Ala Lys Val Asp Ile Pro 385	390	395
75	Lys Asp Thr Leu Ile Thr Tyr Asp Met Val Asn Leu Asp Glu Asp His 405	410	415

ES 2 372 942 A1

Ile Ile Thr Lys Leu Arg Lys Arg Gln Asp Glu Leu Gly Leu
420 425 430

5

<210> 51

<211> 1317

<212> DNA

10 <213> *Brucella abortus*

<400> 51

```
15  atgacaacaa atgtggcact ggtcgggctg gcgcgcgatc tgcgagcgcg ggcggaacc      60
    ggcaaaccga tccgcatcgg cttatcggc gcaggcgaaa tgggcacgga tatcgttaca      120
20  cagggtggcac ggatgcaggg cattgaggtc ggcgcccttt ccgcccgccg cctgcccac      180
    actttcaagg ccatccgcac cgcctatggc gacgaagaaa acgcgcgcga agccacgacc      240
25  gaatccgcaa tgaccgcgc catcgaagcg ggcaagattg ccgttactga cgataacgac      300
    ctgatcctct ccaaccgct catcgatgtc atcatcgacg cgacgggcat tccggaagtg      360
    ggcgccgaga caggcattgc agccatccgc aatggcaagc accttgtcat gatgaatgtc      420
30  gaggccgatg tcaccattgg cccatatctt aaagcgcagg ccgacaagca gggcgtgatt      480
    tattcgctgg gcgcgggtga tgagccgtca tcctgcatgg aactgatcga gtttgtttcg      540
35  gcccttggct atgaagtggc ttcggctggc aagggcaaga acaaccgct caatttcgac      600
    gccacaccgc acgattacag gcaagaagca gaccgccgga acatgaatgt gcgcctgctg      660
40  gtcgagttta tcgatggctc caaaaccatg gtggagatgg cggcgattgc caatgccacc      720
    ggccttgtgc cggacattgc cggcatgcac ggccccgcg ccagcattga ccagctaagc      780
45  cacacgctga tcccacaggc cgaaggcggc gttctcagca aaagcggcgt ggttgattat      840
    tccatcggca agggcgtttc gccggcgctg ttcgtcgtgg cgaagatgga ccacccccgc      900
    ctcaacgaac gtctggaaga cctgaaaatc ggcaaaggcc cttatttcac cttccaccgg      960
50  ccatatcatc tgacttcaact tgaagtgccg ctgaccgctc cccgcgtcgt gcttcacggc      1020
    aagactgaca tggtgccgct gccaaagccg gtggcggaag tctgcgcggt tgccaaaaag      1080
55  gacatgcagc ccggcgagca tctggacgct atcggccaat attgctaccg ctcatggatc      1140
    atgacgggtc cggaagcgcg cgccgcgaag gccattcctt gcggcctgct ccagaacggc      1200
60  acggtcatcg cgccgatcaa aaaaggcgaa ctcatcacct atgccaatgc cgccccgcag      1260
    cccggctcca ggatcgcgga actgcgcgct ttgcaggatg ctatgtagg gcaatag      1317
```

65



OFICINA ESPAÑOLA
DE PATENTES Y MARCAS

ESPAÑA

②① N.º solicitud: 201031068

②② Fecha de presentación de la solicitud: 14.07.2010

③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤① Int. Cl.: Ver Hoja Adicional

DOCUMENTOS RELEVANTES

Categoría	Documentos citados	Reivindicaciones afectadas
X	CHAIN PSG et al. Whole-Genome Analyses of Speciation Events in Pathogenic Brucellae. Infection and Immunity. Diciembre 2005. Vol. 73 (12), páginas 8353–8361, todo el documento.	1-4
A	WO 02083720 A2 (BACHER A. & ROHDICH F.) 24.10.2002, páginas 1-5; reivindicaciones 1, 5, 10, 12, 15, 16.	1-8
A	WO 0111055 A1 (BACHER A. & ZENK M.) 15.02.2001, página 2, línea 25 – página 3, línea 18; reivindicaciones 1,2,12,14-17.	1-8
A	WO 0211673 A2 (SMITHKLINE BEECHAM CORPORATION) 14.02.2002, página 1, línea 29 – página 2, línea 30; página 4, línea 28 – página 6, línea 8; página 8, líneas 10-15.	1-8
A	ROHMER M et al. Isoprenoid biosynthesis as a novel target for antibacterial and antiparasitic drugs. Current Opinion in Investigational Drugs. 2004. Vol. 5(2), páginas: 154-162, todo el documento.	1-8

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

El presente informe ha sido realizado

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe
28.11.2011

Examinador
M. D. García Grávalos

Página
1/5

CLASIFICACIÓN OBJETO DE LA SOLICITUD

C07K14/195 (2006.01)

C12N9/90 (2006.01)

C12N15/61 (2006.01)

C12P23/00 (2006.01)

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C07K, C12N, C12P

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, MEDLINE, NPL, XPESP, EMBASE, BIOSIS, GOOGLE SCHOLAR, EBI.

Fecha de Realización de la Opinión Escrita: 28.11.2011

Declaración

Novedad (Art. 6.1 LP 11/1986)	Reivindicaciones 5-8	SI
	Reivindicaciones 1-4	NO
Actividad inventiva (Art. 8.1 LP11/1986)	Reivindicaciones 5-8	SI
	Reivindicaciones 1-4	NO

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

Base de la Opinión.-

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

1. Documentos considerados.-

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	CHAIN PSG et al. Infection and Immunity. Diciembre 2005. Vol. 73(12), páginas 8353–8361.	2005
D02	WO 02083720 A2	24.10.2002
D03	WO 0111055 A1	15.02.2001
D04	WO 0211673 A2	14.02.2002
D05	ROHMER M et al. Current Opinion in Investigational Drugs. 2004. Vol. 5(2), páginas: 154-162.	2004

2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración

La presente solicitud divulga una enzima, aislada de la bacteria *Brucella abortus*, cepa 2308, que corresponde a la secuencia de aminoácidos definida como SEQ ID NO:41 (reivindicaciones 1-4), así como su uso para la síntesis de isoprenoides, particularmente de 2-C-metil-D-eritritol 4-fosfato (MEP) (reivindicaciones 5 y 6). La invención se refiere también a un método de obtención de MEP (reivindicaciones 7 y 8).

El documento D01 divulga la secuencia completa del genoma de *Brucella abortus*, cepa 2308, y un análisis de dicha secuencia comparándola con las correspondientes a los genomas de otras especies de *Brucella*, también patógenas para humanos, como *B. melitensis*, cepa 16M, *B. suis*, cepa 1330 y de *B. abortus*, cepa 9-941 (ver todo el documento).

El documento D02 divulga una proteína con actividad enzimática y su participación en la síntesis de isoprenoides. Se refiere también a la secuencia de nucleótidos que la codifica, así como a un vector que contiene dicha secuencia y a su expresión en un cultivo celular (ver páginas 1-5; reivindicaciones 1, 5, 10, 12, 15, 16).

El documento D03 divulga una enzima, así como el gen que la codifica, relacionada con la ruta alternativa (MEP) de biosíntesis de isoprenoides. Se refiere también a un vector que contiene dicha secuencia de nucleótidos y a su expresión en un cultivo celular (ver página 2, línea 25 - página 3, línea 18; reivindicaciones 1, 2, 12, 14-17).

El documento D04 divulga un péptido y el polinucleótido que lo codifica, aislado de la bacteria *Haemofilus influenza*, relacionado con la familia de enzimas reductoisomerasas. Se refiere también a un método para modular la actividad de esta enzima e interferir en la ruta alternativa de biosíntesis de isoprenoides (ver página 1, línea 29 - página 2, línea 30; página 4, línea 28 - página 6, línea 8; página 8, líneas 10-15).

El documento D05 divulga la ruta alternativa de biosíntesis de isoprenoides como una nueva diana terapéutica para *screening* de productos antibacterianos y antiparasitarios. Se refiere especialmente a la enzima DXP reductoisomerasa, a su papel en esta ruta de síntesis de isoprenoides y sugiere la búsqueda de otras enzimas con esta actividad (ver todo el documento).

1. NOVEDAD (Art. 6.1 y 8.1 LP 11/1986)**1.1. REIVINDICACIONES 1-8**

La presente solicitud divulga una enzima, aislada de *Brucella abortus*, cepa 2308, que corresponde a la secuencia de aminoácidos definida como SEQ ID NO:41; y su uso para la síntesis de isoprenoides, particularmente de 2-C-metil-D-eritritol 4-fosfato (MEP).

El documento D01 se considera el más cercano al Estado de la Técnica ya que anticipa una enzima, aislada de *Brucella abortus*, cepa 2308, que corresponde a la secuencia de aminoácidos definida en la presente solicitud como SEQ ID NO:41. Esta secuencia de aminoácidos, así como la de nucleótidos que la codifica, se encuentran en la base de datos del NCBI con referencias *Swiss-Prot* Q2YIM3 para aminoácidos y BAB2_0264 para nucleótidos.

Según lo divulgado en el documento D01, las reivindicaciones 1-4 no cumplen con el requisito de novedad (Art. 6.1 LP 11/1986).

Las reivindicaciones 5-8 cumplen con el requisito de novedad (Art. 6.1 LP 11/1986)

2. ACTIVIDAD INVENTIVA (8.1 LP 11/1986)**2.1. REIVINDICACIONES 5-8**

La diferencia entre el documento D01 y la presente invención radica en el uso de la enzima reivindicada para la síntesis de isoprenoides, particularmente de 2-C-metil-D-eritritol 4-fosfato (MEP). Aunque los documentos D02 y D03, anticipan enzimas relacionadas con la ruta alternativa de biosíntesis de isoprenoides, estas enzimas no se corresponden con la secuencia reivindicada, ni han sido aisladas de *Brucella abortus*.

En consecuencia, la invención tal y como se recoge en las reivindicaciones 5-8 cumple con el requisito de actividad inventiva (Art. 8.1 LP 11/1986).

Los documentos D04 y D05, se refieren al estado de la técnica y no se consideran relevantes en relación con el objeto de la invención.